

---

# **TECHNICKÁ UNIVERZITA V LIBERCI**

Fakulta mechatroniky a mezioborových inženýrských studií

Studijní program: N2612 – Elektrotechnika a informatika

Studijní obor: 3906T001 – Mechatronika

## **Rozpoznávání lidských emocí na základě pořízeného obrazu obličeje**

## **Emotion Recognition from Human Face**

### **Diplomová práce**

Autor: Bc. **Lukáš Rosůlek**

Vedoucí práce: Ing. Josef Chaloupka, Ph.D

Konzultant: Ing. Zbyněk Koldovský, Ph.D

**V Liberci 28.5.2009**

## **Prohlášení**

Byl(a) jsem seznámen(a) s tím, že na mou diplomovou práci se plně vztahuje zákon č. 121/2000 Sb. o právu autorském, zejména § 60 (školní dílo).

Beru na vědomí, že Technická univerzita v Liberci (TUL) má právo na uzavření licenční smlouvy o užití mé diplomové práce a prohlašuji, že **souhlasím** s případným užitím mé diplomové práce (prodej, zapůjčení, atd.).

Jsem si vědom(a) toho, že užít své diplomové práce či poskytnout licenci k jejímu využití mohu jen se souhlasem TUL, která má právo ode mne požadovat přiměřený příspěvek na úhradu nákladů, vynaložených univerzitou na vytvoření díla (až do jejich skutečné výše).

Diplomovou práci jsem vypracoval(a) samostatně s použitím uvedené literatury a na základě konzultací s vedoucím diplomové práce a konzultantem.

Datum:

Podpis:

## **Poděkování**

Na tomto místě bych rád poděkoval vedoucímu své diplomové práce panu Ing. Josefu Chaloupkovi, Ph.D. a panu Ing. Zbyňku Koldovskému, Ph.D za jejich cenné rady a připomínky. Rád bych také poděkoval celému kolektivu Fakulty mechatroniky TUL za příjemné studijní prostředí.

## **Abstrakt**

Tato diplomová práce se zabývá problematikou rozpoznávání emocí na základě pořízeného obrazu obličeje. Práce je rozdělena tématicky na čtyři části, přičemž v první části je uveden teoretický přehled problematiky rozpoznávání, kde jsou vysvětleny základní pojmy spojené s rozpoznáváním emocí a obecný postup klasifikace.

Druhá část obsahuje řešerši na dané téma, kde jsou prezentovány jednotlivé přístupy pro rozpoznávání emocí. Zde jsou představeny čtyři odlišné studie využívané pro klasifikaci, která rozhodne o identitě dané emoce. Třetí část se zabývá principem jednotlivých metod, kde je vysvětlen postup při návrhu. Poslední část tvoří experimenty s cílem porovnat vybrané naprogramované algoritmy.

**Klíčová slova:** emoce, strukturální model, analýza hlavních komponent PCA, analýza významných komponent RCA, lineární diskriminační analýza LDA, Fischerova lineární diskriminační analýza FLDA

## **Abstract**

This Diploma thesis deals with the problem of recognizing human emotions from taken photographs. The thesis is divided into four sections. The first section covers base theoretical knowledge about face recognition. The basic concepts of emotion recognition are stated as well as the general technique of classification.

The second part contains the recherche, where are presented individual approaches for emotion recognition. There are introduced four alternative methods of classification, that resolve the identity of current emotion. The third part deals with principles of each method, where is explained the procedure of design. The final part contains experiments, where chosen algorithms are programmed and compared with each others.

**Keywords:** emotion, structural model, Principal Component Analysis PCA, Relevant component analysis RCA, Linear discriminant analysis LDA, Fisher's linear discriminant analysis FLDA

## Obsah

<b>Úvod</b>	<b>8</b>
<b>1. Přehled základních pojmů</b>	<b>9</b>
1.1 Klasifikace emocí.....	9
1.2 Klasifikace rozpoznávacích algoritmů.....	12
1.2.1 Statistické modely.....	12
1.2.2 Strukturální modely.....	12
1.2.3 Hybridní model.....	12
1.3 Transformační prostor.....	13
<b>2. Přístupy pro rozpoznávání emocí</b>	<b>15</b>
2.1 Rozpoznávání emocí pomocí PCA a RBFN.....	15
2.1.1 Vyjmutí specifických rysů.....	15
2.1.2 Implementace přístupu k rozpoznávání emocí.....	16
2.1.3 Výsledky klasifikace.....	17
2.2 Rozpoznávání emocí pomocí Elmanovy neuronové sítě.....	19
2.2.1 Vyjmutí specifických rysů.....	19
2.2.2 Implementace přístupu k rozpoznávání emocí.....	21
2.2.3 Výsledky klasifikace.....	21
2.3 Rozpoznávání emocí pomocí SVM.....	22
2.3.1 Vyjmutí specifických rysů.....	22
2.3.2 Implementace přístupu k rozpoznávání emocí.....	24
2.3.3 Výsledky klasifikace.....	24
2.4 Rozpoznávání emocí pomocí FLDA a RCA.....	25
2.4.1 Vyjmutí specifických rysů.....	25
2.4.2 Implementace přístupu k rozpoznávání emocí.....	27
2.4.3 Výsledky klasifikace.....	27
<b>3. Neuronové sítě</b>	<b>29</b>
3.1 Výkonný prvek umělé neuronové sítě.....	29
3.2 Pracovní fáze umělé neuronové sítě.....	30
3.2.1 Adaptivní fáze.....	30
3.2.2 Aktivní fáze.....	31
3.3 RBFN síť.....	32
3.4 Elmanova síť.....	33
<b>4. Analýza komponent</b>	<b>35</b>
4.1 Analýza hlavních komponent – PCA.....	35
4.1.1 Trénování metodou PCA.....	37
4.1.2 Rozpoznávání metodou PCA.....	39
4.2 Individuální přístup analýzy hlavních komponent – IPCA.....	39
4.3 Analýza významných komponent – RCA.....	40

<b>5. Diskriminační analýza</b>	<b>43</b>
5.1 Lineární diskriminační analýza – LDA.....	43
5.1.1 Trénování lineární diskriminační metody.....	43
5.1.2 Nalezení vektoru $\mu$ .....	44
5.1.3 Nalezení matice $\Sigma$ .....	45
5.1.4 Rozpoznávání lineární diskriminační funkcí.....	45
5.2 Fischerova lineární diskriminační analýza – FLDA..	46
5.2.1 Trénování metodou FLDA.....	47
5.2.2 Rozpoznávání metodou FLDA.....	48
<b>6. Měření vzdálenosti vektorů</b>	<b>49</b>
6.1 Manhattan vzdálenost. ....	49
6.2 Euklidova vzdálenost.. ....	49
6.3 Mahalanobisova vzdálenost.....	49
<b>7. Strojové učení – SVM</b>	<b>50</b>
7.1 Optimální separující nadrovina pro separabilní data .....	51
7.2 Optimální separující nadrovina pro neseparabilní data.....	55
7.3 Nelineární SVM.....	56
<b>8. Experimenty</b>	<b>58</b>
8.1 Popis výrazu obličeje.. ....	58
8.2 Výsledky klasifikace... ....	60
8.2.1 Výsledky metody PCA.. ....	60
8.2.2 Výsledky metody RCA.. ....	60
8.2.3 Výsledky metody IPCA. ....	61
8.2.4 Výsledky metody FLDA.....	61
8.2.5 Výsledky metody LDA.. ....	62
8.2.6 Výsledky metody FLDA+RCA. ....	62
<b>Závěr</b>	<b>63</b>
<b>Použitá literatura</b>	<b>64</b>

## Úvod

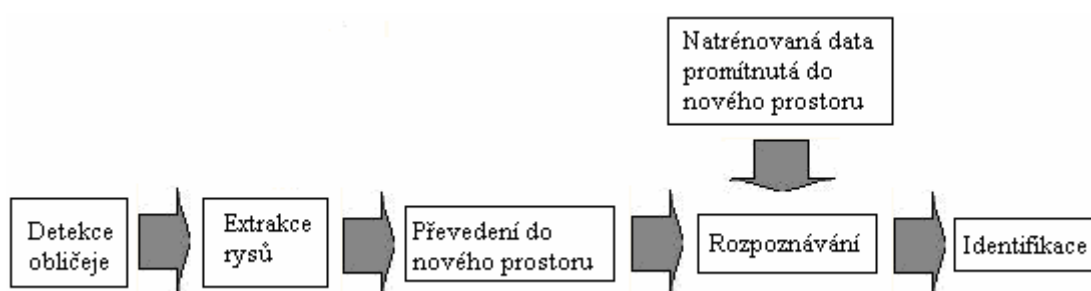
V posledních letech došlo k výraznému rozvoji technik rozpoznávání obličeje. Díky tomu se stále více výzkumníků z oblastí biometrie a počítačového vidění zabývá problematikou identifikace obličeje získaného pomocí snímací techniky, ať už ve formě obrázků nebo videosekvencí. Tento obecný zájem mezi výzkumníky, kteří pracují v různých oblastech informačních technologií, byl především motivován lidskou schopností rozpoznávat lidi a jejich emoce. Již v minulosti se lidé zabývali popisem a zkoumáním jednotlivých emočních výrazů. V roce 1971 Paul Ekman a S. Friesen vydali publikaci popisující výraz obličeje pro šest základních emocí (radost, smutek, překvapení, strach, znechucení a hněv). Kromě mluveného slova využíváme při komunikaci i výraz naší tváře vyjadřující náš vnitřní stav. Dle studie Alberta Mehrabiana [25] je při mezilidské komunikaci výraz v obličeji významnější než informace obsažená v promluvě, protože stejná informace vyslovená s různým výrazem ve tváři může mít naprosto jiný význam.

Některé emoce záměrně schováváme a snažíme se potlačit jejich projevy, proto určitá část emotivního prožívání vždy zůstane skryta. Schopnost rozpoznat emoce roste s tím, jak dobře daného člověka známe a chápeme situaci, která dané emoce vyvolala. U některých zaměstnání je potřeba rozpoznávat emoce, a to zejména u psychicky náročných profesí. Podstatnou nevýhodou počítačů oproti lidem je nedokonalost pochopení situace, která příslušné emoce vyvolala.

Typický systém pro rozpoznávání emocí z výrazu se skládá ze dvou částí. První část se zabývá popisem výrazu obličeje, druhá část jeho klasifikací (rozpoznáváním). Při strukturálním rozpoznávání se ještě před klasifikací provádí vyextrahování charakteristických rysů obličeje, což není u statického rozpoznávání potřeba. Těmito rysy mohou být např. poloha očí, špičky nosu či úst. Často se však v rámci zjednodušení a snížení redundance dat neuchovávají přesné pozice jednotlivých rysů obličeje, nýbrž pouze jejich relativní poloha jako např. vzdálenost mezi jedním okem a obočím, šířka a výška úst, apod. Mnoho metod využívá možnosti převodu obrázku do vektorové podoby a jeho zobrazení do určitého prostoru, kde je řešena podobnost vůči referenční množině. Existuje celá řada vhodných metod a postupů využívaných při rozpoznávání emocí, záleží ale zejména na způsobu použití a kvalitě zpracovávaných dat.

## 1. Přehled základních pojmů

Úkolem rozpoznávacího systému je určit identitu dané emoce nalézající se v získaném obraze. Nejprve musí být obličej ve vstupním obraze nebo videosekvenci lokalizován, tzn. nalezneme jeho přesnou pozici. Dalším krokem je předzpracování, kde se provádí normalizace velikosti nebo jasových hodnot. Posledním krokem, jenž je předmětem této práce, je rozpoznávání již detekovaného obličeje s danou emocí. Tento úkol lze ještě rozdělit na dvě části, kterými jsou nalezení prostoru, jenž bude vhodnou reprezentací dané emoce a klasifikace, která provede rozhodnutí o identitě.



Obr.1.1: Proces rozpoznávání.

### 1.1 Klasifikace emocí

Podněty, které mají pro člověka nějaký význam, vyvolávají emoce. Ty se utvářely v průběhu evoluce a mají nezanedbatelný význam z hlediska adaptace, protože řídí jednání a reakce člověka. Umožňují nám rychle reagovat na nepředvídatelné události, aniž bychom vědomě museli vyhodnocovat situaci. Poskytují informace o vnitřním stavu druhého člověka a o příčinách, které jednotlivé emoce vyvolaly.

Emoce, které neovlivňujeme svými myšlenkami a příliš se nemění během života, nazýváme primární vrozené emoce. Ty jsou základním mechanismem pro reakce člověka. Mezi primární emoce patří např. smutek, strach či radost. Naopak sekundární emoce vznikají pomaleji, rozvíjí se až několik měsíců po narození. Můžeme je do jisté míry ovlivňovat myšlením a jsou vytvářeny především zhodnocením situace, např. oddanost, hrdost, vina a stud.

Pojem základní emoce může nabývat tří odlišných významů. V prvním případě se jednotlivé základní emoce od sebe fundamentálně liší a nemohou být od sebe odvozeny. Podle Paula Ekmana jsou univerzální napříč všemi kulturami a jsou tedy evolučně podmíněny. Již zmíněný Paul Ekman, bývalý profesor psychiatrie na Kalifornské univerzitě, objevil, že tyto základní emoce jsou spjaty s mimickými výrazy tváře [28].



Objevené „mikrovýrazy“, malé mimické pohyby, které trvají méně, než čtvrtinu sekundy, se objevují na naší tváři nezávisle na naší vůli a významně přispívají k určení příslušných emocí. Výraz ovšem není s emocemi pevně spjat, je možné jej záměrně potlačit, nebo naopak předstírat emoce, které neprožíváme.



**Obr.1.2:** Základní emoce podle Paula Ekmana.\*

Druhé pojetí základních emocí předpokládá, že tyto emoce plní adaptivní roli základních životních potřeb. Podle Nekonečného lze identifikovat čtyři základní existenciální situace: zisk, ztráta, hrozba a překážka, jimž odpovídají čtyři základní emoce: radost, smutek, strach a hněv. Autoři zastávající třetí pojetí popisují emoce ve smyslu základních stavebních bloků, ze kterých lze skládat další emoce. Největším zastáncem této teorie byl Robert Plutchik [10]. Přišel s teorií, že existuje několik základních emocí, které nejsou kulturně podmíněny, avšak jsou univerzální pro člověka jako biologický druh. Jeho model obsahuje osm základních emocí: přijetí, zlost, očekávání, znechucení, radost, strach, smutek, překvapení.

---

\* [25]

Autor	Základní emoce
James	strach, žal, vztek, láska
McDougall	strach, znechucení, euforie, podřízenost, něha, pochyby
Watson	strach, láska, zlost
Arnold	hněv, odpor, odvaha, sklíčenost, touha, zoufalství, strach, nenávist, naděje, láska, smutek
Mowrer	bolest, potěšení
Izard	hněv, opovržení, znechucení, úzkost, strach, vina, zájem, radost, hanba, překvapení
Plutchik	příjetí, zlost, očekávání, znechucení, radost, strach, smutek, překvapení
Gray	vztek, zděšení, úzkost, radost
Ekman, Friesen	hněv, odpor, strach, radost, smutek, překvapení
Panksepp	očekávání, strach, vztek, panika
Tomkins	zlost, opovržení, zájem, překvapení, hanba, radost, znechucení, úzkost, strach
Fridja	touha, utrpení, pochybnost, překvapení, zájem, štěstí
Wiener, Graham	štěstí, smutek
Oatley, Johnson-Laird	hněv, odpor, úzkost, radost, smutek

Tab.1.1: Základní emoce – Přehled teorií.

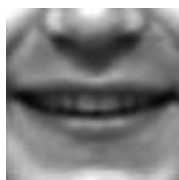
Jak je zřejmé z tabulky (1.1), autoři se neshodují nejen ve výčtu jednotlivých základních emocí, ale ani v jejich počtu. Neshody ve výčtu jednotlivých emocí pramení z faktu, že autoři mají odlišné definice toho, co vůbec emoce jsou a co už za emoce pokládat nelze.

## 1.2 Klasifikace rozpoznávacích algoritmů

Rozpoznávání lidské tváře je natolik zajímavý problém, že se jím zabývají výzkumníci po celém světě už řadu let, a proto existuje velké množství literatury. Díky této pozornosti vzniklo mnoho metod automatického rozpoznávání. Nejzajímavější rozdělení jsou založeny na schopnosti člověka rozpoznat emoce podle celkového vzezření nebo schopnosti rozpoznání na základě nějakých lokálních charakteristických rysů, např. podle očí, nosu nebo úst. Podle uvedeného způsobu můžeme klasifikovat tři skupiny algoritmů [24].

### 1.2.1 Statistické modely

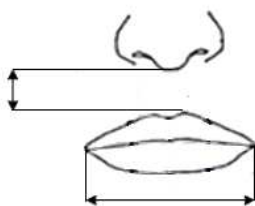
V těchto modelech jsou kombinovány tvar obličeje, nosu, očí a úst společně s barvou kůže. Je nutné, aby byl obličej vždy na fotografii v definované pozici.



Obr.1.3: Statistický model.

### 1.2.2 Strukturální modely

Tyto modely jsou založeny na analýze lokálních vlastností obličeje, tzn. např. vzdálenost očí, výška a šířka nosu, atd.



Obr.1.4: Strukturální model.

### 1.2.3 Hybridní modely

Tyto modely kombinují vlastnosti předešlých dvou, přičemž svým charakterem nejlépe odpovídají lidskému způsobu rozpoznávání.

### 1.3 Transformační prostor

Obrázky se převádějí do transformačního prostoru, který redukuje dimenzi dat a usnadňuje tak následnou klasifikaci. Transformaci lze provádět dvěma způsoby:

1. Extrakce charakteristických rysů: Vytvoření podmnožiny nových rysů kombinací rysů z originálního obrázku.

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \rightarrow \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} = f \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}$$

2. Výběr charakteristických rysů: K dalšímu zpracování se vybere taková podmnožina rysů nesoucí nejvíce informace.

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \rightarrow \begin{pmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{iNM} \end{pmatrix}$$

Problém extrakce charakteristických rysů je definován tak, že je snaha nalézt takové zobrazení  $y = f(x) : R^N \rightarrow R^M, M < N$ , kde vektor  $y_i \in R^M$  si zachovává maximum informace původních charakteristických rysů  $x_i \in R^N$ . Zobrazení  $y = f(x)$  je ale obecně nelineární. Výběr charakteristických rysů tedy bude záviset na konkrétním problému, a proto se často celý problém zjednodušuje omezením pouze na lineární zobrazení  $y = Wx$ , kde  $W$  představuje nový prostor, do kterého jsou charakteristické rysy  $x$  promítnuty.

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \rightarrow \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{pmatrix} = \begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1N} \\ w_{21} & w_{22} & \cdots & w_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{M1} & w_{M2} & \cdots & w_{MN} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}$$

Mezi zástupce této skupiny patří analýza hlavních komponent PCA a lineární diskriminační analýza LDA. Získaný obličejový prostor by měl co nejlépe zachycovat charakteristické rysy, které jsou u jednotlivých emocí různé a naopak potlačit společné. Příliš mnoho informací může vést ke zhoršení výsledků rozpoznávání, protože tyto

informace mohou zachycovat změny v osvětlení nebo nežádoucí šum. Změny v osvětlení obrázků mohou mít největší vliv na výslednou klasifikaci. Tento problém lze vypořádat z obrázku (1.5), zatímco obrázky patří stejné osobě, vlivem osvětlení není tato skutečnost patrná [24].



**Obr.1.5:** Rozdílné osvětlení obrázku.\*

---

\* [24]

## 2. Přístupy pro rozpoznávání emocí

V následujících statích jsou popsány studie užívané pro automatické rozpoznávání emocí na základě pořízeného obrazu obličeje.

### 2.1 Rozpoznávání emocí pomocí PCA a RBFN

V této studii se rozpoznávání emocí provádí pomocí analýzy hlavních komponent PCA a radiálních základních funkcí neuronových sítí RBFN. Autorem této práce je *Daw-Tung Lin* z univerzity v Taipei [4]. Předmětem rozpoznávání je sedm emočních výrazů (radost, smutek, překvapení, strach, znechucení, hněv a neutrální výraz).

#### 2.1.1 Vyjmutí specifických rysů

Obrázky obličeje jsou posuzovány globálně bez zaměření na jednotlivé specifické rysy. Obrázek (2.1) znázorňuje první přístup, kde autor využívá celý obličej pro následnou klasifikaci. Dalším přístupem je nevyužívat celý obličej, ale vzít v úvahu pouze oblast okolo očí a úst (viz. Obr.2.2). Tento přístup se využívá, protože z očí, obočí a rtů můžeme významně rozlišovat změny jednotlivých výrazů.



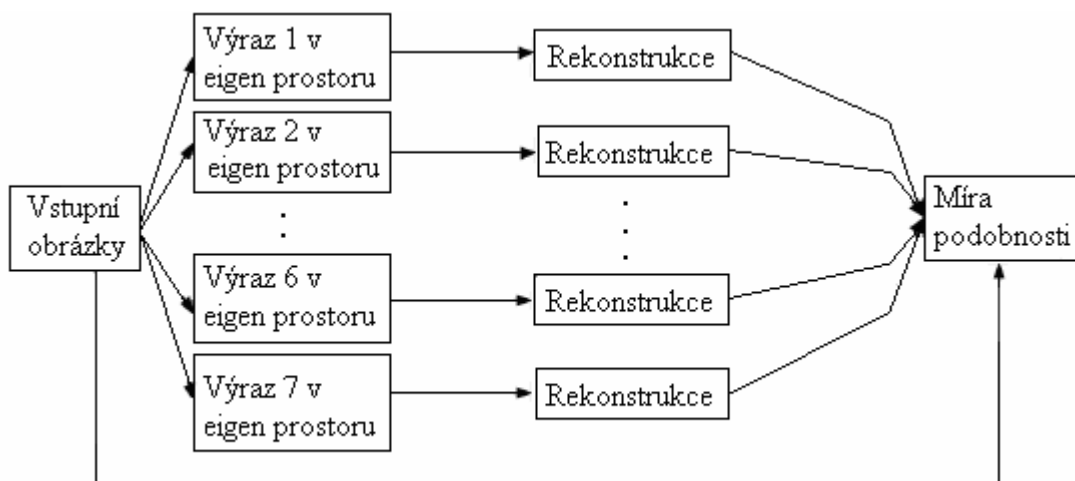
Obr.2.1: Výrazy celého obličeje.\*



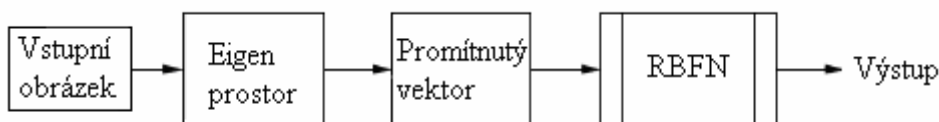
Obr.2.2\*

\* Obrázky jsou převzaty ze studie, jejímž autorem je Daw-Tung Lin z univerzity v Taipei. [4]

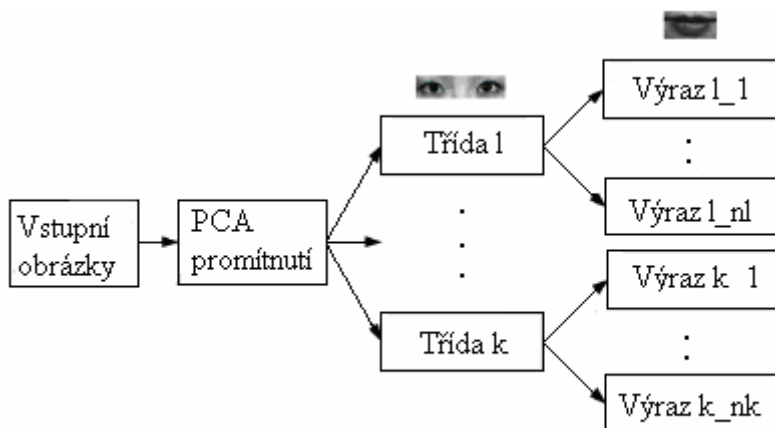
## 2.1.2 Implementace přístupu k rozpoznávání emocí



Obr.2.3: Klasifikační procedura PCA.



Obr.2.4: Klasifikační blokové schéma jednovrstvového RBFN.



Obr.2.5: Blokové schéma hierarchického RBFN.

Pro rozpoznávání výrazů autor využívá kromě PCA a RBFN i hierarchický model (HRBFN). V HRBFN se klasifikační proces dělí na dvě části. V první vrstvě jsou výrazy rozčleněny do  $k$  klasifikací podle vlastních\* rysů z oblasti očí (vlastní rysy jsou výstupem metody PCA). Cílem každé klasifikace je tedy odlišit od sebe jednotlivé třídy výrazů

---

\* eigen

z dané skupiny emocí (úsměv, překvapení, atd.). Počet výrazů, které mají být rozpoznávány v druhé části, je  $n_k$  pro danou třídu  $k$  (viz. Obr.2.5).

### 2.1.3 Výsledky klasifikace

PCA	Úspěšnost [%]	PCA+RBFN	Úspěšnost [%]
Obličej	89,21	Obličej	86,33
Ústa	83,45	Ústa	92,09
Oči	87,77	Oči	82,73

**Tab.2.1:** Procentuální úspěšnost klasifikace.

Metoda	Smutek	Radost	Znechucení	Neutrální	Překvapení	Strach	Hněv
PCA	42,86	68,42	70	80	85	89,47	80
PCA+RBFN	28,57	52,63	25	50	30	100	30

**Tab.2.2:** Procentuální úspěšnost klasifikace z oblasti očí (pro každý výraz zvlášť).

Metoda	Smutek	Radost	Znechucení	Neutrální	Překvapení	Strach	Hněv
PCA	61,9	78,95	50	75	85	78,9	55
PCA+RBFN	61,9	68,42	70	60	75	78,95	75

**Tab.2.3:** Procentuální úspěšnost klasifikace z oblasti úst (pro každý výraz zvlášť).

RBFN	Smutek	Radost	Znechucení	Neutrální	Překvapení	Strach	Hněv
Smutek	<b>6</b>	1	1	0	0	13	0
Radost	0	<b>10</b>	1	0	0	8	0
Znechucení	1	1	<b>15</b>	0	0	13	0
Neutrální	0	3	0	<b>10</b>	0	7	0
Překvapení	0	0	0	3	<b>6</b>	14	0
Strach	0	0	0	0	0	<b>16</b>	0
Hněv	0	0	0	0	0	14	<b>6</b>

**Tab.2.4:** Klasifikační tabulka pro RBFN klasifikátor z oblasti očí. Tato tabulka znázorňuje, do jaké třídy výrazů je zařazen vyšetřovaný výraz (např. smutek je zařazen 13krát do výrazu strach).\*

\* Ostatní klasifikační tabulky též znázorňují, do jaké třídy výrazů je zařazen vyšetřovaný výraz.



RBFN	Smutek	Radost	Znechucení	Neutrální	Překvapení	Strach	Hněv
Smutek	13	0	1	1	0	0	6
Radost	4	13	1	1	0	0	0
Znechucení	4	2	14	0	0	1	1
Neutrální	3	0	0	12	0	4	1
Překvapení	3	1	1	0	15	0	0
Strach	1	0	1	1	1	15	0
Hněv	3	1	1	0	0	0	15

Tab.2.5: Klasifikační tabulka pro RBFN klasifikátor z oblasti úst.

Z klasifikačních tabulek je patrné, že do třídy výrazu strach nebo smutek bylo často přiřazeno hodně výrazů z jiných tříd (hněv, radost, atd.). S ohledem na tyto nedostatky při klasifikaci autor zamýšlí rozdělení klasifikačního procesu do různých skupin a obrázky z oblasti úst jsou trénovány jednotlivě pro každý klasifikátor. Byly navrženy tři postupy (případy). První je tvořen dvěma podmnožinami, jedna obsahuje výraz překvapení a druhá všechny ostatní výrazy (A). Druhý tvoří také dvě podmnožiny, jedna obsahuje neutrální výraz a překvapení, druhá ostatní výrazy (B). Třetí opět tvoří dvě podmnožiny, jedna obsahuje výrazy úsměv, překvapení, neutrální výraz a druhá všechny ostatní (C). Klasifikační výsledky navrženého HRBFN jsou uvedeny v tabulce (2.6). Nejvyšší dosažené rozpoznávací skóre z oblasti očí je 95,68%, z oblasti úst 72,66%.

	Případ A		Případ B		Případ C	
	oči	ústa	oči	ústa	oči	ústa
trénovaná sada	100	100	100	100	100	100
testovací sada	95,68	71,94	93,52	72,66	91,37	69,78

Tab.2.6: Procentuální úspěšnost klasifikace HRBFN modelu.

Případ	část	Smutek	Radost	Znechucení	Neutrální	Překvapení	Strach	Hněv
A	oči	100	100	100	95	80	94,73	100
	ústa	66,67	73,68	70	65	80	73,68	75
B	oči	100	100	100	70	90	100	95
	ústa	57,14	68,42	80	70	85	78,95	70
C	oči	95,24	84,21	90	95	90	84,21	100
	ústa	57,14	57,89	60	95	65	73,68	80

Tab.2.7: Procentuální úspěšnost klasifikace HRBFN modelu pro jednotlivé výrazy.

## 2.2 Rozpoznávání emocí pomocí Elmanovy neuronové sítě

V této studii se rozpoznávání emocí provádí pomocí Elmanovy neuronové sítě. Autory této práce jsou *Shen-Chuan Tai*, *Hung-Fu Juany*, *Kuo-Chen Chung*, *Yu-Yi Liao* a *Chien-Shiang Hong* z Cheng Kung univerzity v Thajsku [17]. Předmětem rozpoznávání je šest emočních výrazů (radost, smutek, překvapení, strach, znechucení, hněv).

### 2.2.1 Výjmutí specifických rysů

Pro rozpoznávání je třeba vybrat hlavní rysy (body) reprezentující jednotlivé emoční výrazy. V prvním obrázku je tedy ručně označeno sedmnáct specifických bodů.

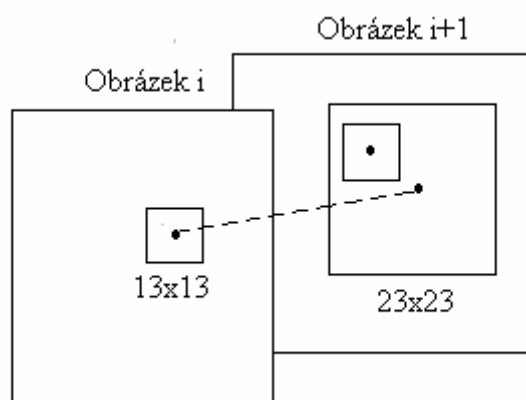


**Obr.2.6:** Obrázek ukazuje, jaké body jsou využívány pro další zpracování.\*

Každý bod je středem okna o velikosti 13x13. Pro automatické sledování bodů v sekvenčním obraze se využívá vzájemné korelace založené na metodě optického toku. V prvním kroku musíme nejprve vypočítat vzájemnou korelaci z okna 13x13 v prvním obraze s 23x23 oknem v druhém obraze. Pozice s maximální korelací je odhadnuta jako pozice hlavního bodu v druhém obraze. Každý hlavní bod je vypočítán odečtením standardní pozice v prvním obraze od aktuální standardní pozice, přičemž pozice všech hlavních bodů je normalizována vůči špičce nosu.

\* Myšlenka tohoto způsobu vyjmutí specifických rysů, ovšem s částečnou modifikací, je využívána i v této práci.[17]

Obrázek (2.7) znázorňuje implementaci této metody na dvou po sobě jdoucích obrázcích.



**Obr.2.7:** Výpočet vzájemné korelace.

Jednotlivé emoční výrazy lze popsat pomocí matematického modelu. Tento matematický model vychází z obrázku (2.6).

$$\text{Výška očí} \quad ((Y5-Y6)+(Y9-Y10))/2 \quad (2.1)$$

$$\text{Šířka očí} \quad ((X4-X3)+(X8-X7))/2 \quad (2.2)$$

$$\text{Vzdálenost obočí od duhovky} \quad ((Y1-Y11)+(Y2-Y12))/2 \quad (2.3)$$

$$\text{Šířka úst} \quad X15-X14 \quad (2.4)$$

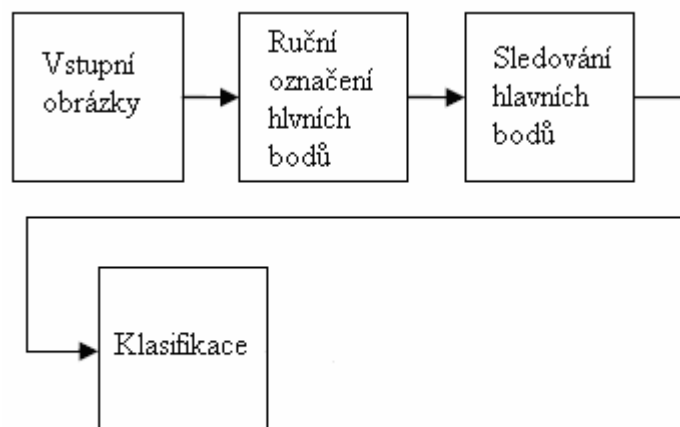
$$\text{Výška úst} \quad Y16-Y17 \quad (2.5)$$

$$\text{Prohloubenina horního rtu} \quad Y13-Y16 \quad (2.6)$$

$$\text{Vzdálenost oka od líce} \quad ((Y11-Y13)+(Y12-Y13))/2 \quad (2.7)$$

Výrazy v těchto rovnicích představují souřadnice bodů ve dvourozměrném prostoru.

## 2.2.2 Implementace přístupu k rozpoznávání emocí



**Obr.2.8:** Klasifikační blokové schéma.

## 2.2.3 Výsledky klasifikace

Emoce	Úspěšnost [%]
Radost	100
Překvapení	96,4
Strach	82,2
Hněv	89,5
Znechucení	96,1
Smutek	87,6
<b>Průměr</b>	<b>92</b>

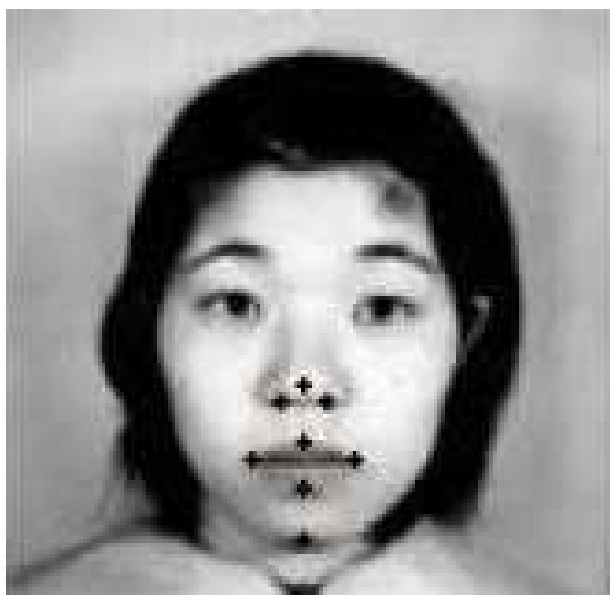
**Tab.2.8:** Výsledky klasifikace.

## 2.3 Rozpoznávání emocí pomocí SVM

V této studii se rozpoznávání emocí provádí pomocí podpůrných vektorů SVM. Autorem této práce je *Porawat Visutsak* ze státní vývojové instituce v Thajsku [27]. Předmětem rozpoznávání je šest emočních výrazů (hněv, odpor, strach, štěstí, smutek a překvapení). Využívá se souboru hlavních bodů (rysů) vyjmutých z pořízeného obrazu obličeje a uvažuje se jejich pohyb pro každý emoční výraz. Vektor posunutí ze všech pořízených bodů vstupuje do SVM klasifikátoru.

### 2.3.1 Vyjmutí specifických rysů

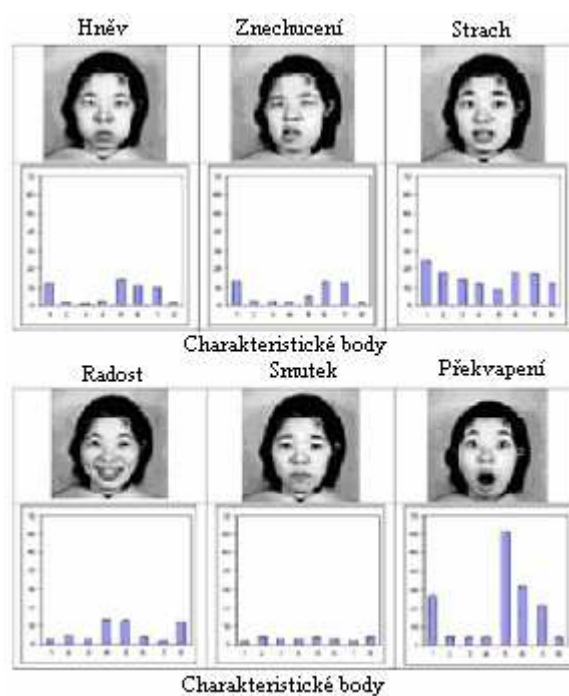
Body reprezentující jednotlivé výrazy jsou na obrázku označeny ručně. Je zjištěno posunutí mezi neutrálním výrazem obličeje a šesti emočními výrazy (hněv, odpor, strach, štěstí, smutek, překvapení). Tímto je stanoven charakteristický pohybový vzor pro každý emoční výraz.










- |                     |                 |
|---------------------|-----------------|
| 1.hrot nosu         | 5.pravý roh úst |
| 2.pravá nosní dírka | 6.levý roh úst  |
| 3.levá nosní dírka  | 7.spodní ret    |
| 4.horní ret         | 8.brada         |

**Obr.2.9:** Vyjmutí specifických rysů.\*

\* Obrázek je převzat ze studie, jejímž autorem je *Porawat Visutsak* ze státní vývojové instituce v Thajsku. [27]



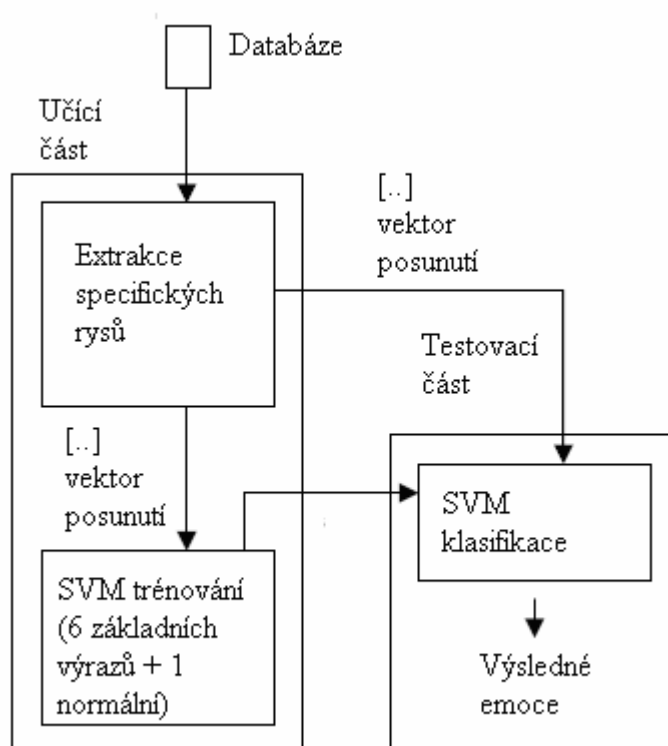
**Obr.2.10:** Posunutí mezi normálním výrazem a ostatními emočními výrazy.\*

	Výrazy	Rad	Sm	Př	Zl	Zn	Str
Neutrální		3,03	2,45	1,74	2,00	1,90	1,77
Radost		4,84	1,42	1,35	1,23	1,23	1,23
Smutek		1,45	4,61	1,52	1,55	2,13	2,39
Překvapení		2,35	1,81	4,87	1,81	1,71	2,29
Zlost		1,19	2,16	1,55	4,74	2,84	1,39
Znechucení		1,35	2,90	2,45	2,65	4,81	3,32
Strach		1,23	4,19	3,74	2,16	3,42	4,26

**Obr.2.11:** Ohodnocení pro každý ze šesti emočních výrazů (5-vysoká podobnost, 1-nízká podobnost). Toto ohodnocení se využívá při vyhodnocení každého výrazu.

\* [27]

### 2.3.2 Implementace přístupu k rozpoznávání emocí



Obr.2.12: Blokové schéma SVM klasifikace.

### 2.3.3 Výsledky klasifikace

SVM	Úspěšnost [%]
Radost	91,5
Překvapení	97,5
Strach	66,7
Hněv	67,7
Znechucení	62,3
Smutek	61
<b>Průměr</b>	<b>74,5</b>

Tab.2.9: Výsledky klasifikace.

## 2.4 Rozpoznávání emocí pomocí FLDA a RCA

V této studii se emoce rozpoznávají pomocí Fischerovy lineární diskriminační analýzy FLDA a analýzy významných komponent RCA. Výsledky navrženého klasifikátoru jsou porovnány s výsledky SVM klasifikace. Autory této práce jsou *M. Sorci, G. Antonini* a *Jean-Philippe Thiran* z institutu zpracování signálu ve Švýcarsku [18]. Předmětem rozpoznávání je sedm emočních výrazů (hněv, odpor, strach, štěstí, smutek, překvapení a neutrální výraz).

### 2.4.1 Vyjmutí specifických rysů

V tomto případě se využívá pro vyjmutí specifických rysů AAM modelu (Active Appearance Model), kde je obličej reprezentován 55 body. Aktivní „vzhledový“ model je takový model, který neobsahuje jen informaci o tvaru, ale je v něm zahrnuta i informace o jasů jednotlivých bodů modelu. Tato metoda tedy vychází z toho, že objekt nacházející se v neznámém obraze má specifický tvar, kterému odpovídá i specifický jas nebo barva příslušných bodů. Pro každý obraz je definována množina bodů  $x_i = (x_{i,1}, y_{i,1}, \dots, x_{i,n}, y_{i,n})$ , kde  $n$  je počet bodů. V prvním kroku musíme nejprve transformovat všechny obrázky tak, aby měly objekty stejnou orientaci, rotaci a měřítko. Dále je vypočítána střední hodnota  $\bar{x}$  z vektorů  $x_i$  a kovarianční matice  $S$  (dimenze  $2N \times 2N$ ) odchylek  $x_i$  od průměrného tvaru  $\bar{x}$ .

$$\bar{x} = \frac{1}{s} \sum_{i=1}^s x_i \quad (2.8)$$

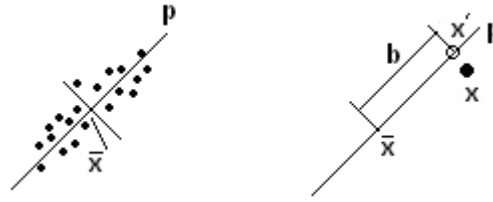
$$S = \frac{1}{s-1} \sum_{i=1}^s (x_i - \bar{x})(x_i - \bar{x})^T \quad (2.9)$$

Je zřejmé, že některé body mohou více přispívat ke změně tvaru, proto je použito principu analýzy hlavních komponent PCA (kapitola 4.1), která zredukuje počet proměnných jen na ty nejdůležitější podle toho, jak hodně ovlivňují tvar objektu. K výpočtu nového tvaru podle rovnice (2.10), kde  $b_s$  udává odchylku od středního tvaru, je využita matice  $P_s = (p_1, p_2, \dots, p_s)$ , která představuje vlastní vektory kovarianční matice  $S$  seřazené sestupně.

$$x = \bar{x} + P_s b_s \quad (2.10)$$



Zmenšováním a zvětšováním vektoru  $b_s$  můžeme měnit tvar modelu. Příпустné hodnoty parametrů  $b_i$  se pohybují v intervalu  $\langle -3\sqrt{\lambda_i}, 3\sqrt{\lambda_i} \rangle$ , kde  $\lambda_i$  jsou vlastní čísla kovarianční matice  $S$  seřazená sestupně.



**Obr.2.13:** Každý bod může být promítnut na přímku  $p$ ,  $x \approx x' = \bar{x} + bp$ , kde  $b$  je vzdálenost od bodu  $x$  ke střední hodnotě  $\bar{x}$ .

Díky tomu, že jsou vektory  $p_i$  seřazené podle důležitosti, je zajištěno, že při vynechání posledních vektorů dojde k nejmenší chybě. Abychom mohli modely porovnávat podle jasů, musíme každý objekt transformovat do středního tvaru vypočítaného z trénované množiny pomocí označených bodů. Protože může mít scéna obrázků různé osvětlení, je nutné provést normalizaci jasů. Jeden ze vzorků zvolíme jako referenční. Jeho jas je tedy vybrán jako základní a následně je pomocí iterativní metody normalizována hodnota jasů vždy vůči nově vypočtené střední hodnotě. Na upravená data je opět aplikována metoda PCA, čímž je získána matice  $P_g$  a vektor  $b_g$ . Dosazením do vzorce (2.11) a změnou  $b_g$  můžeme vytvořit nové jasové hodnoty pro tvar daný  $\bar{x}$ . Tímto postupem je získán lineární model:

$$g = \bar{g} + P_g b_g. \quad (2.11)$$

Následně vypočítáme vektor  $b$ , abychom mohli měnit jak tvar, tak i hodnotu jasů modelu. Tento vektor lze formulovat vztahem:

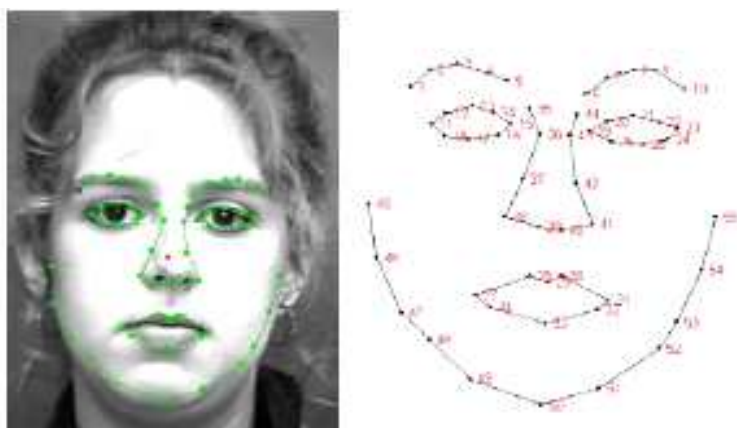
$$b = \begin{pmatrix} W_s b_s \\ b_g \end{pmatrix} = \begin{pmatrix} W_s P_s^T (x - \bar{x}) \\ P_g^T (g - \bar{g}) \end{pmatrix}, \quad (2.12)$$

kde  $W_s$  je diagonální matice vah. Pro každý prvek trénované množiny je získán vektor  $b$  a opětovnou aplikací metody PCA na tyto vektory je získána matice  $Q$  a vektor  $c$ . Jejich dosazením do rovnice (2.14) můžeme generovat nové tvary a jim odpovídající jasy modelu.

$$b = Qc \quad (2.13)$$

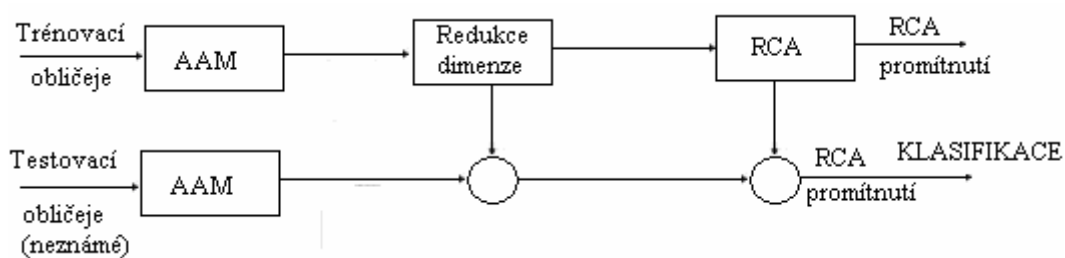
$$x = \bar{x} + P_s W_s Q_s c, \quad g = \bar{g} + P_g Q_g c, \quad Q = \begin{pmatrix} Q_s \\ Q_g \end{pmatrix} \quad (2.14)$$

Nový model je vytvořen tak, že jsou vypočítány jasy  $g$  pro tvář  $\bar{x}$  a tento obraz je následně transformován do tvaru daného  $x$  [3], [18], [22].



Obr.2.14: Kontura obličeje reprezentována 55 body.\*

## 2.4.2 Implementace přístupu k rozpoznávání emocí



Obr.2.15: Blokové schéma klasifikace.

## 2.4.3 Výsledná klasifikace

FLDA+RCA	Úspěšnost [%]	SVM	Úspěšnost [%]
Radost	100	Radost	100
Překvapení	95	Překvapení	85
Strach	72,72	Strach	90,91
Hněv	47,06	Hněv	64,7
Znechucení	100	Znechucení	100
Smutek	94,12	Smutek	64,7
Neutrální	93,33	Neutrální	93,33
<b>Průměr</b>	<b>86</b>	<b>Průměr</b>	<b>85,5</b>

Tab.2.10: Výsledky klasifikace.

\*[18]

Metoda	Úspěšnost [%]
LDA	85,217
GDA+RCA	82,609
GDA	82,609
RCA	76,522

**Tab.2.11:** Klasifikační poměr pro další testované metody (GDA – zobecněná nelineární diskriminační analýza).

<b>FLD+RCA</b>	Radost	Překvapení	Strach	Hněv	Znechucení	Smutek	Neutrální
Radost	<b>18</b>	0	0	0	0	0	0
Překvapení	0	<b>19</b>	0	0	1	0	0
Strach	1	0	<b>8</b>	0	1	0	1
Hněv	0	0	0	<b>8</b>	7	0	2
Znechucení	0	0	0	0	<b>17</b>	0	0
Smutek	0	0	0	1	0	<b>16</b>	0
Neutrální	0	0	0	1	0	0	<b>14</b>

**Tab.2.12:** Klasifikační tabulka pro FLDA+RCA klasifikátor.

<b>SVM</b>	Radost	Překvapení	Strach	Hněv	Znechucení	Smutek	Neutrální
Radost	<b>18</b>	0	0	0	0	0	0
Překvapení	0	<b>17</b>	0	0	1	1	1
Strach	0	0	<b>10</b>	0	1	0	0
Hněv	2	0	0	<b>11</b>	3	0	1
Znechucení	0	0	0	0	<b>17</b>	0	0
Smutek	0	0	0	6	0	<b>11</b>	0
Neutrální	0	0	0	1	0	0	<b>14</b>

**Tab.2.13:** Klasifikační tabulka pro SVM klasifikátor.

Z výsledné klasifikace je patrné, že metoda FLDA+RCA dává srovnatelné výsledky jako SVM. V [18] autoři upřednostňují FLDA+RCA před SVM, protože se v této klasifikaci nemusí ladit a nastavovat žádné parametry. Nastavení parametrů je nutné v SVM klasifikaci, kde je výsledek dosažen pomocí zdlouhavého a subjektivního ladění klasifikátoru.

### 3. Neuronové sítě

Neuronová síť (Neutral network) je jedním z výpočetních modelů používaných v umělé inteligenci. Umělá neuronová síť se skládá z umělých neuronů, jejichž předobrazem je biologický neuron. Neurony si navzájem předávají signály, které jsou transformovány pomocí přenosových funkcí. Za umělou neuronovou síť můžeme považovat takovou strukturu pro distribuované paralelní zpracování dat, která se skládá z obvykle velmi vysokého počtu vzájemně propojených výkonných prvků. Každý z těchto prvků může přijímat libovolný konečný počet dat a na další výkonné prvky může předávat libovolný konečný počet informací o stavu svého velmi rozvětveného výstupu. Každý výkonný prvek transformuje vstupní data na výstupní podle přenosové funkce. Funkci umělé neuronové sítě můžeme chápat jako transformaci  $T$  vstupního signálu  $X$  na výstupní signál:

$$\bar{Y} = T(\bar{X}). \quad (3.1)$$

Funkci  $T$  lze implementovat jako transformační funkci neuronové sítě, která má nejméně tři vrstvy s dopřednou vzájemnou vazbou [5].

#### 3.1 Výkonný prvek umělé neuronové sítě

Výkonným prvkem umělé neuronové sítě je formální neuron, jehož funkci lze matematicky popsat podle vztahu:

$$y = S\left(\sum_{i=1}^N w_i x_i + \Theta\right), \quad (3.2)$$

kde

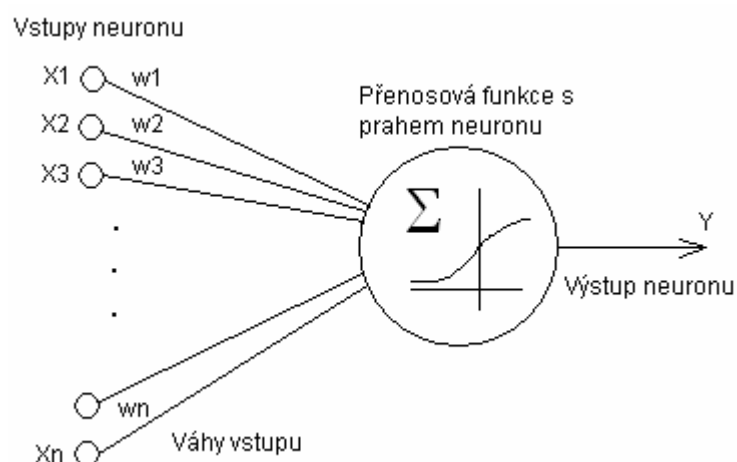
$x_i$  jsou vstupy neuronu, těchto vstupů je celkem  $N$ ,

$w_i$  jsou synaptické váhy,

$S$  je obvykle nelineární přenosová funkce neuronu,

$\Theta$  je práh.

Prahová hodnota znamená bariéru, kterou musí vstupní signál neuronu překonat, aby se mohl dál šířit neuronovou sítí. Hodnota prahu tedy určuje, kdy je neuron aktivní resp. pasivní.



Obr.3.1: Základní model jednoduchého neuronu.

### 3.2 Pracovní fáze umělé neuronové sítě

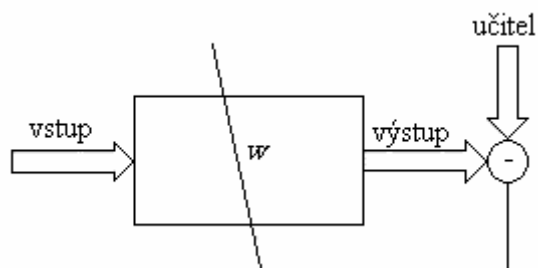
Rozlišujeme 2 fáze – adaptivní, kde se síť učí a aktivní, kde vykonává naučenou činnost – vybavuje si. Paměť je nejčastěji reprezentována hodnotami váhových koeficientů jednotlivých vstupních neuronů [5].

#### 3.2.1 Adaptivní fáze

Charakteristickou vlastností umělých neuronových sítí je schopnost učení. To probíhá řízeným nastavováním vazeb mezi neurony, jež je realizováno změnou synoptických vah neuronů. Tyto váhy se nastavují tak, aby síť reagovala na vstupní vzory odpovídajícími vzory výstupními.

#### Učení s učitelem

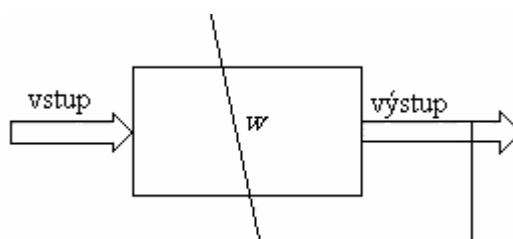
Učení s učitelem také někdy nazýváme chybové učení. Jedná se o přístup, kdy je do sítě poslána dvojice hodnot. Jedna hodnota je vstupní a druhá výstupní zadaná učitelem. Reálný výstup sítě, který vznikne jako odezva sítě na daný vstup, se potom porovnává s výstupem požadovaným a podle velikosti odchylky mezi těmito hodnotami se upravují synoptické váhy spojů sítě.



Obr.3.2: Blokové schéma učení s učitelem.

### Učení bez učitele

Tento přístup je založen na schopnosti neuronových sítí hledat ve vstupech podobné vlastnosti a třídit pak vstupy podle těchto vlastností. Podobné vektory se potom sdružují do tzv. *shluků*. Učící algoritmus nezná hodnoty výstupů, proto se tento princip učení používá právě v případech, kdy neznáme výstupní hodnoty. Principem učení je výpočet vzdáleností mezi vzory se společnými vlastnostmi (shluky) a aktuálními hodnotami vstupu.



Obr.3.3: Blokové schéma učení bez učitele.

### 3.2.2 Aktivní fáze

V druhé fázi neuronová síť reaguje na předložené vstupy změnou hodnot na výstupech, váhy již nejsou upravovány. Na základě vstupu dat do sítě vznikne ve vstupní vrstvě nerovnovážný stav. Zapamatované hodnoty se začnou působením ostatních neuronů přes spoje měnit (aktualizovat), mění se tak dlouho, dokud opět nenastane stabilní rovnovážný stav. Na výstupu se potom nachází požadovaná odezva sítě na tento vstup [5].

### 3.3 RBFN síť

RBFN síť má jednu vstupní, jednu skrytou a jednu výstupní vrstvu. Neurony ve skryté vrstvě jsou radiálního typu, tj. počítá se vzdálenost vstupního vektoru od center shluků. Centra shluků mají vlastnosti dané průměrnými hodnotami všech vzorů příslušné skupiny vstupních dat. Vztah pro aktivační funkci lze definovat jako:

$$\varphi(\vec{x}, \vec{c}_j) = \exp\left(-\frac{1}{2\sigma^2} \|\vec{x} - \vec{c}_j\|^2\right), \quad (3.3)$$

kde

$\vec{x}$  jsou vstupní data,

$\vec{c}_j$  jsou centra shluků ze vstupních dat,

$\sigma$  je prostorový parametr.

Aktivační funkce skrytých neuronů má nejčastěji charakter Gaussovy křivky, přičemž potom platí, že vzory, které jsou daleko od centra shluků, mají nižší vliv na chování sítě. Váhy v první vrstvě (mezi vstupní a skrytou vrstvou) se nastavují na začátku učení a jsou pevné po celou dobu učení. Hodnoty vektorů vah vedoucích k jednotlivým skrytým neuronům jsou reprezentovány centry shluků ve vstupních datech. Váhy v druhé vrstvě (mezi skrytou a výstupní vrstvou) se nastavují „gradientními“ algoritmy, nebo se mohou určit pomocí metody nejmenších čtverců. Výstup neuronové sítě lze popsat vztahem:

$$F_k(x) = \sum_{j=1}^n w_{jk} \cdot \varphi(\vec{x}, \vec{c}_j), \quad (3.4)$$

kde  $w$  jsou synoptické váhy. Úlohu můžeme definovat tak, že máme k dispozici soubor vstupních dat  $\vec{x}_i \in R^p$ ,  $i = 1, 2, \dots, N$ , kterým odpovídají cílové hodnoty  $\vec{d}_i \in R^k$ ,  $i = 1, 2, \dots, N$ . Naším cílem je nalézt takovou funkci  $F : R^p \rightarrow R^k$ , která bude splňovat podmínku:

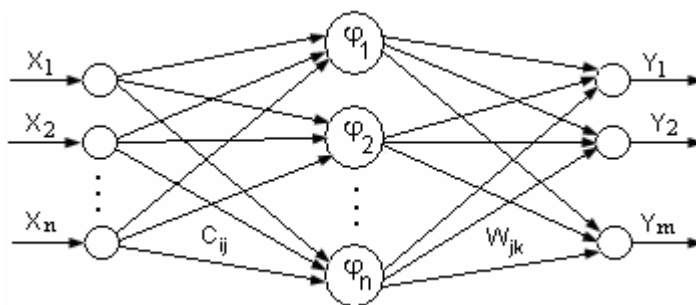
$$F(\vec{x}_i) = \vec{d}_i, \quad i = 1, 2, \dots, N. \quad (3.5)$$

Řešení vychází ze vztahu (3.4). Výslednou rovnici lze formulovat takto:

$$\begin{pmatrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1m} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{N1} & \varphi_{N2} & \cdots & \varphi_{Nm} \end{pmatrix} \begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1k} \\ w_{21} & w_{22} & \cdots & w_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m1} & w_{m2} & \cdots & w_{mk} \end{pmatrix} = \begin{pmatrix} d_{11} & d_{12} & \cdots & d_{1k} \\ d_{21} & d_{22} & \cdots & d_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ d_{N1} & d_{N2} & \cdots & d_{Nk} \end{pmatrix} \quad (3.6)$$

kde  $\varphi_{ij} = \varphi(\bar{x}_i, \bar{c}_j)$ ,  $i = 1, 2, \dots, N$ ,  $j = 1, 2, \dots, m$ . Rovnici (3.6) lze přepsat do zkráceného tvaru  $\varphi \cdot W = D$ . Aby tato rovnice splňovala podmínku (3.5), musíme vypočítat váhovou matici  $W$ , která je získána pomocí metody nejmenších čtverců [4]. Řešením našeho problému je tedy vztah:

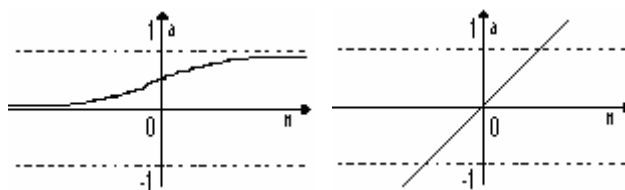
$$W = (\varphi^T \varphi)^{-1} \varphi^T D. \quad (3.7)$$



**Obr.3.4:** Příklad topologie radiální neuronové sítě.

### 3.4 Elmanova síť

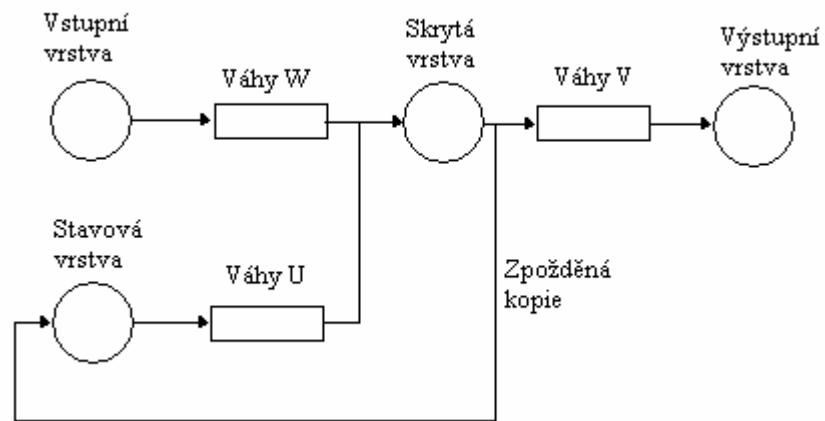
Elmanova síť je architekturou rekurentní umělé neuronové sítě, kde je výstup ze skryté vrstvy kopírován do stavové vrstvy neuronové sítě. Stavová vrstva se pak v dalším časovém kroku stává součástí vstupu. Tato architektura je vhodná zejména při adaptaci sítě. Elmanova síť má tansigmoidální aktivační funkci ve skryté vrstvě a lineární aktivační funkci ve výstupní vrstvě. Tato kombinace se využívá zejména proto, že dvouvrstvá síť může pomocí těchto aktivačních funkcí aproximovat jakoukoli funkci s libovolnou přesností.



**Obr.3.5:** Levá část-tansigmoidální aktivační fce., Pravá část - lineární aktivační fce.

Porovnáním požadovaných výstupních a cílových hodnot je získána posloupnost chyb, která je v každé iteraci využita pro stanovení gradientu chyb pro každý parametr. Získané hodnoty gradientu jsou pak použity pro aktualizaci parametrů sítě. Při trénování rekurentní Elmanovy sítě se využívá algoritmu zpětného šíření, který je přizpůsoben použitému typu rekurentní sítě [21].





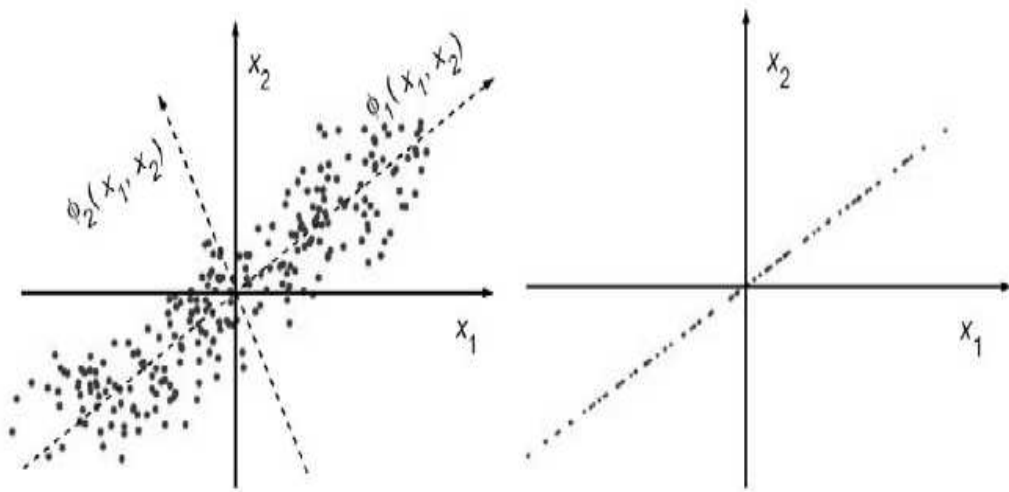
Obr.3.6: Elmanova síť.

## 4. Analýza komponent

### 4.1 Analýza hlavních komponent – PCA

Analýza hlavních komponent (Principal Component Analysis) je metodou redukce dimenze s minimální ztrátou informace v datech, která je založena na transformaci souřadného systému. Dochází k nalezení speciální ortonormální báze prostoru, ve kterém jsou data umístěna. Vektory hledané ortonormální báze jsou uspořádány tak, že první určuje směr obsahující největší možnou informaci a ve směru posledního bázevého vektoru je obsah informace minimální. První hlavní komponenta je lineární kombinací takového rozměru původních dat, který vykazuje největší rozptyl hodnot a každá další  $n$ -tá hlavní komponenta je pak lineární kombinací následujícího rozměru, který představuje nejlépe zachycený rozptyl původních dat, přičemž je vždy ortogonální ke všem  $n-1$  předcházejícím. Metoda PCA je blízká Karhunen-Loèevě transformaci KLT, jenž byla odvozena v souvislosti se zpracováním signálů jako ortogonální transformace s bází  $V=(V_1, V_2, \dots, V_N)^T$ , která pro jakékoliv  $k \leq N$  minimalizuje euklidovskou  $L_2$  rekonstrukční chybu  $e(X)$  pro data  $X$ . Pokud mají data nulový střední vektor, jsou pak formulace PCA a KLT identické [16].

$$e(X) = \left\| X - \sum_{i=1}^k (V_i^T X) \cdot V_i \right\| \quad (4.1)$$



**Obr.4.1:** Redukce dimenze metodou PCA, kde osa  $\phi_1$  znázorňuje směr, ve kterém je rozptyl dat maximální.\*

\* [16]

Myšlenka metody PCA je dobře zachycena na obrázku (4.1). Pokud proložíme osu mezi body vícerozměrného prostoru, můžeme na ni promítnout všechny uvažované body, které jsou nyní rozptýleny kolem ní. Vznikají tak nové body  $X'$ , mezi kterými můžeme nyní zjistit vzdálenosti jednotlivých bodů. Jsou-li na přímce body shromážděny blízko u sebe, bude pak variance nízká, v opačném případě bude vysoká.

**Analýza hlavních komponent je realizována následujícím postupem:**

1. Uspořádáme data do matice  $X$ .
2. Vypočítáme průměrný vektor dat (průměrný řádek matice  $X$ ).
3. Vypočítáme kovarianční matici  $K$ .
4. Spočítáme vlastní čísla a vlastní vektory matice  $K$  (vlastní čísla uspořádaná v absolutní hodnotě od největšího k nejmenšímu  $\lambda_n$  a jim příslušné vektory  $v_n$ ).
5. Vybereme  $n$  hlavních komponent  $v_1, \dots, v_n$  a sestavíme transformační matici  $V$  obsahující prvních  $n$  vlastních vektorů kovarianční matice  $K$ .
6. Promítneme původní data do redukovaného prostoru  $V$  generovaného hlavními komponentami.  $P = V^T X$

Matice  $P$  pak obsahuje původní data kolmo promítnutá do afinního podprostoru dimenze  $n$ , který je nejlepší v tom smyslu, že celková chyba  $e(X)$  způsobená promítáním je minimální [1], [20].

#### 4.1.1 Trénování metodou PCA

Trénování metodou PCA probíhá v několika krocích. Bez ztráty na obecnosti budeme předpokládat, že vstupní data budou mít nulový střední vektor, tzn. že jsou nejprve vycentrovaná. Z těchto nových dat je vytvořena kovarianční matice, ze které jsou následně vypočítány vlastní čísla a vlastní vektory. Nechť  $N$  je počet obrázků pro každý emoční výraz z  $C$  emocí v databázi. Každý obrázek může být reprezentován jako vektor o rozměrech  $d \times 1$ , kde  $d$  je počet pixelů obrázku (popřípadě představuje vyjmuté body nebo vzdálenosti). Každý z těchto vektorů je označen  $x_{ij}$ , kde index  $i$  značí číslo emočního výrazu a index  $j$  označuje číslo obrázku  $i$ -tého výrazu. Tyto vektory pak tvoří sloupce  $d \times CN$  matice  $X$ .

$$X = (x_{11}, x_{12}, x_{13}, x_{21}, \dots, x_{CN}) \quad (4.2)$$

Průměrný vektor lze formulovat vztahem:

$$\mu = \frac{1}{NC} \sum_{i=1}^C \sum_{j=1}^N x_{ij} . \quad (4.3)$$

Od každého obrázku  $x_{ij}$  odečteme celkový obrazový průměr  $\mu$ , čímž je získána nová matice  $A$  s rozměry  $d \times NC$ .

$$a_{ij} = x_{ij} - \mu, \quad A = (a_{11}, a_{12}, \dots, a_{21}, a_{22}, \dots, a_{CN}) \quad (4.4)$$

Takto získaná množina vektorů je vstupem do analýzy hlavních komponent, kde dochází k nalezení množiny  $n$  ortonormálních vlastních vektorů  $v_n$ , které nejlépe popisují distribuci vstupních dat. Hledané vektory jsou hlavními komponentami kovarianční matice  $K = AA^T$ . Řešíme tedy rovnici:

$$AA^T v_n = \lambda_n v_n, \quad n = 1, 2, \dots, d. \quad (4.5)$$

Vlastní čísla  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$  jsou nezáporná a v klesajícím pořadí. Matice  $K$  je řádu  $d \times d$  a u statického modelu rozpoznávání (obrázek reprezentován pixely) je zjišťování  $d$  vlastních čísel výpočetně velmi náročné. Například při použité velikosti obrázků  $64 \times 64$  pixelů je rozměr  $K$   $4096 \times 4096$ . Proto se jejich výpočet provádí pomocí lineární kombinace vektorů získaných řešením výpočtu vlastních čísel matice  $A^T A$ . Toto lze užít pouze za předpokladu  $CN \ll d$ , tedy že počet obrázků je daleko menší než dimenze prostoru.

Následně jsou vypočítány vlastní čísla  $\lambda_n$  a vlastní vektory  $v_n$   $CN \times CN$  rozměrné matice  $A^T A$ , tzn.

$$A^T A v_n = \lambda_n v_n, \quad n = 1, 2, \dots, CN. \quad (4.6)$$

Vlastní čísla matice  $A^T A$  jsou shodná s vlastními čísly matice  $AA^T$ . Vlastní vektory shodné ovšem nejsou, a proto musí být přepočítány. Vlastní vektor  $v_n$  původní kovarianční matice  $AA^T$  je vypočítán vynásobením matice  $A$  s vypočteným vlastním vektorem  $v_n$  a vydělením jeho normou. U velkých databází se může výpočet výrazně zpomalit i při použití této metody pokud  $CN > d$ .

$$v_n = \frac{A v_n}{\|v_n\|} \quad (4.7)$$

Pro normu vektoru platí:

$$\|v_n\| = \sqrt{\sum v_n v_n^*}, \quad (4.8)$$

pokud jsou vlastní vektory reálné:

$$\|v_n\| = \sqrt{\sum v_n^2}. \quad (4.9)$$

Takto vypočtené vlastní vektory (odpovídající vlastním číslům seřazeným od největšího k nejmenšímu) tvoří ortonormální bázi redukovaného prostoru  $V$  označovaného jako *eigenspace* (vlastní prostor)\*. Počet vektorů v *eigenspace* je možné určit dle hodnot vlastních čísel, malá čísla přispívají málo k popisu odchylek mezi obrázky a příslušné vlastní vektory můžeme tudíž ignorovat.

$$V = (v_1, v_2, \dots, v_{CN}) \quad (4.10)$$

Posledním krokem je promítnutí vycentrovaných trénovaných obrázků do vytvořeného redukovaného prostoru  $V$ .

$$P = V^T A \quad (4.11)$$

Tímto krokem je získána matice  $P$  o rozměrech  $CN \times CN$ , která obsahuje trénovaná data kolmo promítnutá do prostoru *eigenspace*.

---

\* *eigenspace* = vlastní prostor

#### 4.1.2 Rozpoznávání metodou PCA

Rozpoznávání metodou PCA spočívá v měření vektorového rozdílu (zvolenou metrikou) vycentrovaného vstupního testovaného obrázku promítnutého do redukovaného prostoru  $V$  se všemi promítnutými trénovanými obrázky. Od vstupního obrázku  $f$  je tedy nejprve odečten obrazový průměr  $\mu$  trénované množiny a následně je vycentrovaný vektor  $t$  transformován do redukovaného prostoru  $V$ .

$$t = f - \mu, \quad J = V^T t \quad (4.12)$$

Testovaný vektor, představující neznámou emoci, je přiřazen emočnímu výrazu v trénované množině na základě nejmenší hodnoty rozdílu.

#### 4.2 Individuální přístup analýzy hlavních komponent – IPCA

Přístup metody IPCA je založen na vypočtení rekonstrukční chyby vzniklé promítáním testovaného obrázku do charakteristického prostoru dané emoce neboli na rozdílu mezi původním testovacím obrázkem a jeho rekonstrukcí po promítnutí do charakteristického prostoru příslušné emoce. Postup trénování je podobný jako u metody PCA s tím rozdílem, že eigenspace je konstruován pro každý emoční výraz zvlášť [13]. Průměrný vektor pro  $i$ -tou emoci lze formulovat jako:

$$\mu_i = \frac{1}{N} \sum_{j=1}^N x_{ij}, \quad (4.13)$$

kde  $x_{ij}$  představuje  $j$ -tý trénovaný obrázek  $i$ -té emoce. Dalším krokem algoritmu je vycentrování původních dat, kde vycentrovaná matice  $A$  není společná pro všechny emoce, ale je vytvořena pro každou zvlášť.

$$s_{ij} = x_{ij} - \mu_i \quad A_i = (s_{i,1}, s_{i,2}, \dots, s_{i,N}) \quad (4.14)$$

Z těchto nových dat je vytvořena kovarianční matice a z ní jsou následně vypočítány vlastní čísla a vlastní vektory. Hledáme takovou množinu ortonormálních vlastních vektorů  $v_{in}$ , které nejlépe popisují distribuci vstupních dat.

$$AA^T v_{in} = \lambda_{in} v_{in} \quad (4.15)$$

Pokud vstupuje do IPCA obrázky reprezentovaný pixely, využívá se pro urychlení výpočtu přístup popsáný v předcházející části.

Takto vypočtené vlastní vektory tvoří ortonormální bázi redukovaného prostoru  $U_i = (u_{i1}, u_{i2}, \dots, u_{in})$ . Proces rozpoznávání pak probíhá tak, že testovací vektor  $f$  je promítnut do charakteristického prostoru každé emoce, následně zrekonstruován a z rozdílu původního a zrekonstruovaného obrázku je vypočítána rekonstrukční chyba. Promítnutí testovaného vektoru do charakteristického prostoru lze formulovat vztahem:

$$g = U_i^T (f - \mu_i). \quad (4.16)$$

Rekonstruovaný vektor  $k$  lze definovat jako:

$$k = \sum_n^Q g u_{in}, \quad (4.17)$$

kde  $Q$  je počet vlastních vektorů odpovídající  $Q$  největším vlastním číslům kovarianční matice  $i$ -té emoce. Protože vypočtený redukovaný prostor není přesnou reprezentací původních dat, vždy při rekonstrukci vznikne residuum  $e$  (čtvercová odchylka) [13].

$$e = \|(f - \mu_i) - k\|^2 \quad (4.18)$$

Testovaný vektor, představující neznámou emoci, je přiřazen emočnímu výrazu v trénované množině na základě nejmenšího rezidua. Pro zrychlení výpočtů můžeme pro vypočtení rezidua  $e$  použít následující vztah:

$$e = \|(f - \mu_i)\|^2 - \sum_{n=1}^Q g^2. \quad (4.19)$$

### 4.3 Analýza významných komponent – RCA

Analýza významných komponent (Relevant component analysis) je podobná analýze hlavních komponent PCA s tím rozdílem, že u této metody není hledán směr reprezentující největší možnou změnu v datech, ale je transformován datový prostor transformační maticí  $W$ , která přiřadí vysokou váhu „významným dimenzím“ a nízkou váhu „bezvýznamným dimenzím“. Algoritmus je založený na takzvaných *chunklets*\*. Chunklets se skládají ze vzorků, které patří stejné, ačkoli neznámé třídě [7], [18]. Mějme vstupní matici  $X$  (neboli vstupní množinu chunklets), kde  $N$  je počet všech obrázků z trénované množiny v databázi.

---

\* V současné době neexistuje vhodný český ekvivalent.

Každý obrázek může být reprezentován jako vektor o rozměrech  $d \times 1$ . Tyto vektory pak tvoří sloupce  $d \times N$  matice  $X$ .

$$X = (x_1, x_2, x_3, \dots, x_N) \quad (4.20)$$

Je nezbytné, aby vstupní data měla nulový střední vektor, tzn. že jsou nejprve vycentrovaná. Od každého vektoru z matice  $X$  tedy odečteme celkový vektorový průměr  $\mu$ , čímž je získána nová vycentrovaná matice  $A$  o rozměrech  $d \times N$ .

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \quad (4.21)$$

$$a_i = x_i - \mu, \quad A = (a_1, a_2, \dots, a_N) \quad (4.22)$$

V dalším kroku je vypočítána střední hodnota  $\mu_c$  ze všech vzorků obsažených v každém sloupci nově vzniklé matice  $A$ . Každý sloupec (chunklet) matice  $A$  obsahuje  $d$  vzorků  $(a_{c,1}, \dots, a_{c,d_c})$ . Z těchto nových dat je dále vytvořena kovarianční matice  $C$  o rozměrech  $d \times d$ :

$$C = \frac{1}{n} \sum_{c=1}^N \sum_{i=1}^{d_c} (a_{c,i} - \mu_c)(a_{c,i} - \mu_c)^T \quad (4.23)$$

kde  $n = \sum_{c=1}^N d_c$ . Pokud by vstupní data obsahovala velký počet vzorků, výpočet kovarianční matice by byl velmi pomalý, v některých případech i nepřijatelný. Následně jsou takto upravená data promítnuta do prostoru generovaného hlavními komponentami pomocí metody PCA:

$$L = V^T C, \quad (4.24)$$

kde  $V$  jsou vlastní vektory kovarianční matice  $K$  (kapitola 4.1). Dalším krokem algoritmu je vytvoření bělící transformace. Tuto transformaci lze vyjádřit vztahem:

$$W = \Lambda^{-\frac{1}{2}} \Phi^T \Rightarrow \text{diag} \left\{ \frac{1}{\sqrt{\lambda_1}}, \frac{1}{\sqrt{\lambda_2}}, \dots, \frac{1}{\sqrt{\lambda_d}} \right\} \Phi^T, \quad (4.25)$$

kde  $\Phi$  je ortogonální matice vlastních vektorů (odpovídající vlastním číslům matice  $L$ ) a  $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_d)$  je diagonální matice s kladnými vlastními čísly matice  $L$  seřazenými sestupně.



Posledním krokem je promítnutí vycentrovaných trénovaných obrázků do nově vytvořeného transformačního prostoru  $W$ .

$$Y = WA \quad (4.26)$$

Rozpoznávání metodou RCA je obdobné jako u PCA metody. Spočívá v měření vektorového rozdílu (zvolenou metrikou) transformovaného vycentrovaného testovacího obrázku se všemi transformovanými trénovacími obrázky. Od vstupního obrázku  $f$  je tedy nejprve odečten obrazový průměr  $\mu$  trénované množiny a následně je vycentrovaný vektor  $t$  transformován do prostoru  $W$ .

$$t = f - \mu, \quad J = Wt \quad (4.27)$$

Testovaný vektor, představující neznámou emoci, je přiřazen emočnímu výrazu v trénované množině na základě nejmenší hodnoty rozdílu.

## 5. Diskriminační analýza

### 5.1 Lineární diskriminační analýza - LDA

Lineární diskriminační metoda umožňuje rozlišení dvou nebo více odlišných tříd námi zkoumaných bodů ve vektorovém prostoru. Množina zvolených proměnných je určena svou odlišností v různých třídách. Tato kapitola vychází z práce [6]. Řešíme klasifikaci vstupních vektorů  $X$  do  $R$  tříd, indikátory tříd označíme  $\omega_1, \omega_2, \dots, \omega_R$ . Vektor příznaků  $X$  je zařazen do té třídy  $\omega^*$ , pro níž platí:

$$\omega^* = \arg \max_r g_r(X). \quad (5.1)$$

Mějme množinu funkcí:

$$g_r(X) = p(X|\omega_r), \quad (5.2)$$

kde  $p(X|\omega_r)$  je hustota pravděpodobnosti udávající rozložení hodnot příznakových vektorů  $X$ . Diskriminační funkce pak definujeme následujícím vztahem:

$$g_r(X) = \ln p(X|\omega_r). \quad (5.3)$$

#### 5.1.1 Trénování lineární diskriminační metody

Pro správnou klasifikaci příznakových vektorů do tříd  $\omega_1, \omega_2, \dots, \omega_R$  je potřeba znát podmíněné hustoty pravděpodobností  $p(X|\omega_r)$ . Předpokládáme, že hustoty pravděpodobností  $p(X|\omega_r)$  odpovídají normálnímu rozložení. Tento předpoklad umožňuje popsat je rovnicí:

$$p(X|\omega_r) = \frac{1}{\sqrt{(2\pi)^N |\Sigma_r|}} \exp\left(-\frac{1}{2}(X - \mu_r)^T \Sigma_r^{-1}(X - \mu_r)\right), \quad (5.4)$$

kde

$N$  je počet prvků příznakového vektoru  $X$ ,

$\mu_r$  je střední hodnota normálního rozdělení,

$\Sigma_r$  je jeho kovarianční matice.

Pro získání hustot pravděpodobností  $p(X|\omega_r)$  je nutné znát pro každou třídu hodnoty parametrů  $\mu_r$  a  $\Sigma_r$ . Tyto parametry jsou získány na základě statistického rozboru

příznakových vektorů, u nichž známe, ke které klasifikační třídě patří. Předpokládejme, že máme množinu trénovaných dat obsahující  $M$  příznakových vektorů patřících k jedné třídě  $\omega_r$ .

$$p(X_1, X_2, \dots, X_M | q) = \prod_{m=1}^M p(X_m | q) \quad (5.5)$$

Dále musíme nalézt pro danou třídu  $\omega_r$  takový parametr  $q^*$ , který maximalizuje věrohodnost (5.5). Tento odhad je maximálně věrohodným odhadem a je vyjádřen vztahem:

$$q^* = \arg \max_q \prod_{m=1}^M p(X_m | q). \quad (5.6)$$

Dosazením hustoty pravděpodobnosti (5.4) do rovnice pro věrohodnost (5.5) je získána podmíněná hustota pravděpodobnosti trénované množiny:

$$\begin{aligned} p(X_1, X_2, \dots, X_M | \mu, \Sigma) &= \\ &= \prod_{m=1}^M \frac{1}{\sqrt{(2\pi)^N |\Sigma|}} \exp\left(-\frac{1}{2}(X_m - \mu)^T \Sigma^{-1}(X_m - \mu)\right). \end{aligned} \quad (5.7)$$

Celou rovnici zlogaritmujeme a výsledek použijeme pro výpočet  $\mu_r$  a  $\Sigma_r$ .

$$\begin{aligned} \ln p(X_1, X_2, \dots, X_M | \mu, \Sigma) &= \\ &= -\frac{NM}{2} \ln(2\pi) - \frac{M}{2} \ln(\Sigma) - \frac{1}{2} \sum_{m=1}^M ((X_m - \mu)^T \Sigma^{-1}(X_m - \mu)) \end{aligned} \quad (5.8)$$

### 5.1.2 Nalezení vektoru $\mu$

Provedeme parciální derivaci rovnice (5.8) podle vektoru  $\mu$ .

$$\begin{aligned} \frac{\partial \ln p(X_1, X_2, \dots, X_M | \mu, \Sigma)}{\partial \mu} &= \\ &= -\frac{1}{2} \sum_{m=1}^M \frac{\partial ((X_m - \mu)^T \Sigma^{-1}(X_m - \mu))}{\partial \mu} = \Sigma^{-1} \sum_{m=1}^M (X_m - \mu) \end{aligned} \quad (5.9)$$

Výraz (5.9) položíme roven nule:

$$\Sigma^{-1} \sum_{m=1}^M (X_m - \mu) = 0 \quad (5.10)$$

a dostaneme:

$$\mu = \frac{1}{M} \sum_{m=1}^M X_m. \quad (5.11)$$

Tento výraz je maximálně věrohodným odhadem parametru  $\mu$  normálního rozdělení.

### 5.1.3 Nalezení matice $\Sigma$

Při výpočtu matice  $\Sigma$  postupujeme obdobně jako u vektoru  $\mu$ . Vypočteme parciální derivaci věrohodnostní funkce podle parametru  $\Sigma$  a položíme ji rovnou nule.

$$\begin{aligned} \frac{\partial \ln p(X_1, X_1, \dots, X_M | \mu, \Sigma)}{\partial \Sigma} &= \\ &= -\frac{M}{2} \frac{\partial \ln |\Sigma|}{\partial \Sigma} - \frac{1}{2} \sum_{m=1}^M \frac{\partial ((X_m - \mu)^T \Sigma^{-1} (X_m - \mu))}{\partial \Sigma} = \\ &= -\frac{M}{2} \Sigma^{-1} + \frac{1}{2} \Sigma^{-1} \sum_{m=1}^M ((X_m - \mu)(X_m - \mu)^T) \Sigma^{-1} = 0 \end{aligned} \quad (5.12)$$

Poté z rovnice (5.12) vyjádříme parametr, pro nějž je hodnota věrohodnostní funkce maximální.

$$\Sigma = \frac{1}{M} \sum_{m=1}^M ((X_m - \mu)(X_m - \mu)^T) \quad (5.13)$$

### 5.1.4 Rozpoznávání lineární diskriminační funkcí

Dosadíme-li do vztahu (5.4) za  $p(X|\omega_r)$  rovnici (5.5) dostaneme:

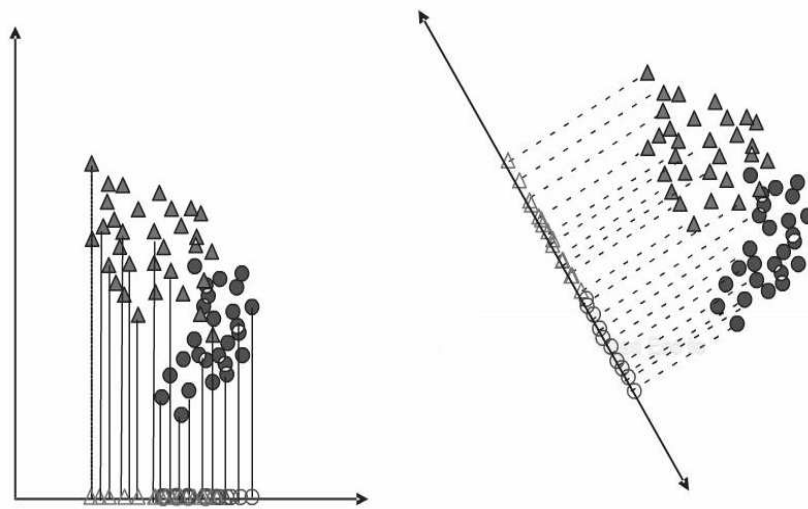
$$g_r(X) = -\frac{N}{2} \ln(2\pi) - \frac{1}{2} \ln |\Sigma_r| - \frac{1}{2} (X - \mu_r)^T \Sigma_r^{-1} (X - \mu_r). \quad (5.14)$$

Rozhodující funkci diskriminačního klasifikátoru lze přepsat do tvaru:

$$g_r(X) = -\ln |\Sigma_r| - (X - \mu_r)^T \Sigma_r^{-1} (X - \mu_r). \quad (5.15)$$

## 5.2 Fischerova lineární diskriminační analýza - FLDA

Mezi nejznámější diskriminační metody patří Fischerova lineární diskriminační analýza FLDA. Tato metoda převádí klasifikační proces do nižší dimenze, kde odstraňuje problém mnohorozměrných dat a minimalizuje ztrátu informace v datech používaných pro následnou klasifikaci. Cílem FLDA je nalézt takovou projekci dat do prostoru nižší dimenze, která umožňuje co nejzřetelnější oddělení jednotlivých tříd. Hledáme tedy takový směr projekce, při kterém budeme schopni na získané přímce co nejlépe lineárně oddělit prvky odlišných tříd.



**Obr.5.1:** Levý obrázek představuje nevhodnou projekci na přímku, při které oddělitelnost tříd není zřejmá. Pravý obrázek naopak znázorňuje optimální projekci zřetelně oddělující obě třídy.\*

**Fischerova lineární diskriminační analýza je realizována následujícím postupem:**

1. Uspořádáme data do matice  $X$ .
2. Vypočítáme průměrný vektor dat přes všechny třídy (průměrný řádek matice  $X$ ) a průměrný vektor z každé třídy zvlášť.
3. Vypočítáme matici mezitřídní variability  $S_B$  a matici vnitrotřídní variability  $S_W$ .
4. Spočítáme vlastní čísla a vlastní vektory matice  $S_W^{-1} S_B$  (vlastní čísla jsou uspořádána v absolutní hodnotě od největšího k nejmenšímu  $\lambda_n$ ).

---

\* [15]

5. Promítneme původní data do nově vytvořeného prostoru  $W_{FLDA}$ .

### 5.2.1 Trénování metodou FLDA

Mějme matici  $X$  (4.2), stejně jako je tomu u metody PCA. Naším cílem je transformovat matici  $X$  tak, aby došlo k těsnějšímu shluknutí objektů patřících do téže třídy a současně k co největšímu oddělení jednotlivých tříd od sebe. Transformaci, která promítne trénovaná data do prostoru nižší dimenze, můžeme popsat rovnicí  $Y = W^T X$ . Náš požadavek na oddělení odlišných tříd odpovídá požadavku, aby v rozptylu  $W^T X$  dominovala složka mezitřídní variability na úkor variability vnitrotřídní. Požadavek na výše popsanou lineární transformaci lze formulovat jako maximalizační úlohu:

$$W_{FLDA} = \arg \max_W \frac{W^T S_B W}{W^T S_W W} = [w_1, w_2, \dots, w_d], \quad (5.16)$$

kde  $S_B$  je definována jako mezitřídní matice odpovídajícího rozptýlení (meztřídní variabilita) a  $S_W$  je rozptylovou maticí konkrétní třídy (vnitrotřídní variabilita). Tyto matice lze popsat rovnicemi:

$$S_B = \sum_{i=1}^C N_i (\mu_i - \mu)(\mu_i - \mu)^T \quad (5.17)$$

$$S_W = \sum_{i=1}^C \sum_{x_k \in X_i} (x_k - \mu)(x_k - \mu)^T \quad (5.18)$$

$$\mu_i = \frac{1}{N} \sum_{j=1}^N x_j, \quad (5.19)$$

kde  $N$  je počet trénovaných obrázků v dané třídě,  $C$  je počet odlišných tříd,  $\mu$  střední hodnota přes všechny třídy a  $\mu_i$  je střední hodnota z dané třídy  $X_i$ . Dále platí vztah:

$$S_T = S_B + S_W, \quad (5.20)$$

kde  $S_T$  je variační maticí vstupních vektorů  $X$ . Tuto variační matici lze formulovat vztahem:

$$S_T = \sum_{i=1}^{CN} (x_i - \mu)(x_i - \mu)^T, \quad (5.21)$$

kde  $CN$  je počet všech trénovaných obrázků v matici  $X$ .

Nechť je dána rovnice:

$$S_B w_n = \lambda_n S_W w_n, \quad n=1,2,\dots,d. \quad (5.22)$$

V následující části se budeme zabírat řešením této rovnice. Matice  $S_W$  je regulární, a tudíž k ní existuje inverzní matice. Hledáme takovou množinu vlastních vektorů  $w_n$ , které nejlépe popisují distribuci vstupních dat.

$$S_W^{-1} S_B w_n = \lambda_n w_n, \quad n=1,2,\dots,d \quad (5.23)$$

Tato rovnice bude splněna, pokud  $\lambda_n$  bude rovno libovolnému vlastnímu číslu matice  $S_W^{-1} S_B$  a vektor  $w_n$  bude vlastní vektor příslušející k  $\lambda_n$ . Vlastní čísla  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$  jsou nezáporná a v klesajícím pořadí. Vlastní vektory  $w_1, w_2, \dots, w_d$  jsou množinou nenulových řešení rovnice (5.23) příslušejících k nenulovým vlastním číslům  $\lambda_1, \lambda_2, \dots, \lambda_d$ . Posledním krokem je promítnutí trénovaných dat do tohoto vytvořeného prostoru  $W_{FLDA}$ .

$$W_{FLDA} = (w_1, w_2, \dots, w_d) \quad (5.24)$$

FLDA tedy transformuje trénovaná data  $X$  do nového prostoru  $W_{FLDA}$ , jehož dimenze může nabývat maximálně velikosti  $C-1$ , kde  $C$  udává počet tříd [1], [15].

$$Y = W_{FLDA}^T \cdot X \quad (5.25)$$

### 5.2.2 Rozpoznávání metodou FLDA

Rozpoznávání metodou FLDA spočívá v měření vektorového rozdílu vstupního testovacího obrázku promítnutého do charakteristického prostoru  $W_{FLDA}$  se všemi rovněž promítnutými trénovanými obrázky.

## 6. Měření vzdálenosti vektorů

K měření rozdílu vektorů se používá tzv. metrika. Pro každý z trénovaných obrázků je vypočten zvolenou metrikou rozdíl a obrázek s nejmenší výslednou hodnotou je určen jako nejpodobnější.

### 6.1 Manhattan vzdálenost

Tato vzdálenost bývá často označována jako city vzdálenost, protože je podobná algoritmu průjezdu kolem severoamerických městských bloků s kolmými ulicemi, od bodu A do bodu B. Jedná se o vzdálenost mezi dvěma body v Euklidově prostoru s pevným kartézským souřadnicovým systémem.

$$D(x, y) = \sum_i |x_i - y_i| \quad (6.1)$$

### 6.2 Euklidova vzdálenost

Euklidova vzdálenost je chápána jako ordinární vzdálenost mezi dvěma body, na kterou můžeme opakovaně aplikovat Pythagorův teorém. Euklidova norma  $D$  měří rozdíl dvou vektorů, přičemž jsou sčítány čtverce odchylek jednotlivých prvků vektorů.

$$D(x, y) = \sum_i (x_i - y_i)^2 \quad (6.2)$$

### 6.3 Mahalanobisova vzdálenost

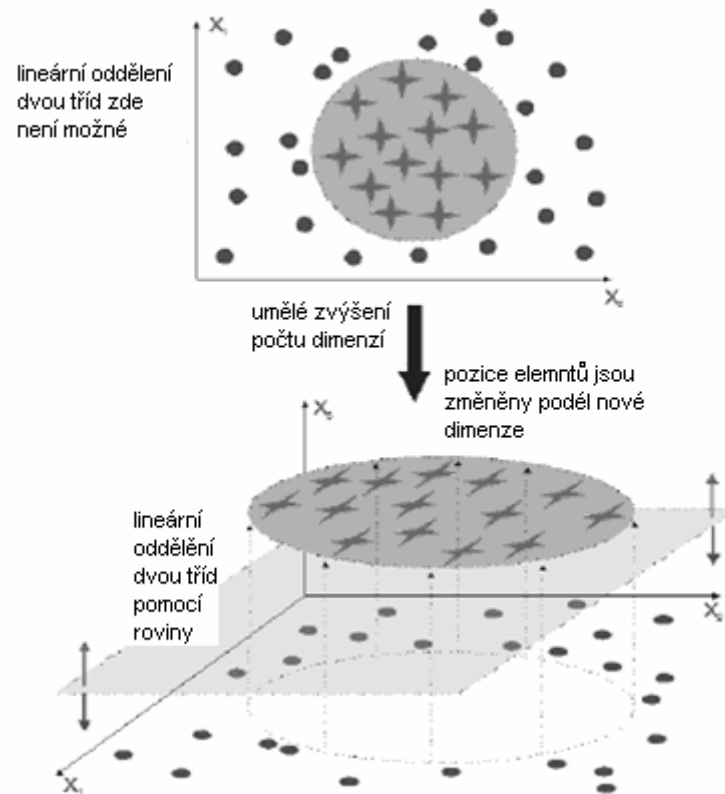
Mahalanobisova vzdálenost je vhodná ke zjišťování rozdílu neznámého vektoru od známé množiny dat. Od Euklidovy normy se liší tím, že je brána v potaz korelace datové množiny a její hodnota se vynásobením některého z vektorů konstantou nemění. Mahalanobický rozdíl vektoru  $x$  od známé množiny dat s průměrem  $\mu$  a kovarianční maticí  $C$  je dán vztahem:

$$d(x) = \sqrt{(x - \mu)^T C^{-1} (x - \mu)}. \quad (6.3)$$



## 7. Podpůrné vektory - SVM

Podpůrné vektory SVM (Support vector machines) tvoří kategorii tzv. jádrových algoritmů (kernel machines). Tyto metody jsou schopny reprezentovat vysoce složité nelineární funkce, které převedou původní vstupní prostor dat do jiného, vícerozměrného, kde lze již od sebe oddělit jednotlivá data lineárně. Obecně platí, že  $N$  datových bodů lze lineárně oddělit v prostoru s  $N-1$  nebo více dimenzemi [19]. Cílem SVM je nalézt takovou přímku či obecně nadrovinu, která bude od sebe oddělovat objekty jednotlivých tříd. Nevýhodou SVM je, že jádra a příslušné parametry je nutné pro konkrétní úlohu volit ručně.



**Obr.7.1:** Princip vzniku možnosti lineárního oddělení dvou tříd s nelineárními hranicemi pomocí přidané dimenze.\*

\* [19]

## 7.1 Optimální separující nadrovina pro separabilní data

K dispozici máme množinu trénovaných dat  $D\{x_i, y_i\}$ ,  $i=1..l$ , kde  $x_i$  jsou vstupní vektory souřadnic v  $d$ -dimenzionálním prostoru  $x_i \in R^d$  a  $y \in \{-1, 1\}$ . Jak už bylo řečeno v předcházející části, cílem lineární SVM učící metody je nalezení optimální nadroviny oddělující od sebe zástupce odlišných tříd. Optimální oddělovací nadrovinu lze definovat jako takovou nadrovinu, okolo které je maximálně široký pás nikoho a na každé straně jsou příklady jedné třídy. Matematicky lze rovnici oddělovací nadroviny zapsat jako:

$$w^T x + b = 0, \quad (7.1)$$

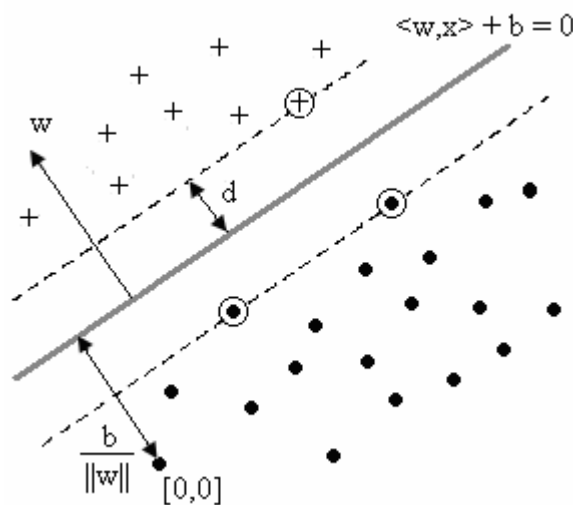
kde  $w$  je normálový vektor (vektor vah) a  $|b|/\|w\|$  je vzdálenost nadroviny k počátku souřadné soustavy (viz. Obr.7.2). Úlohu můžeme formulovat tak, že trénovaná data vyhovují následujícím podmínkám:

$$w^T x_i + b \geq +1 \text{ pro } y_i = +1 \quad (7.2)$$

$$w^T x_i + b \leq -1 \text{ pro } y_i = -1. \quad (7.3)$$

Tyto dvě podmínky lze též formulovat vztahem:

$$y_i \cdot (w^T x_i + b) - 1 \geq 0 \quad \forall i. \quad (7.4)$$



**Obr.7.2:** Lineární rozdělovací nadrovina.

Nyní vezmeme v úvahu body, pro které platí podmínka (7.2). Tyto body jsou odděleny nadrovinou  $H_1$ :

$$w^T x_i + b = 1 \quad (7.5)$$

s normálou  $w$  a vzdáleností od počátku  $|1 - b| / \|w\|$ . Podobně platí, že body splňující podmínku (7.3) jsou odděleny nadrovinou  $H_2$ :

$$w^T x_i + b = -1 \quad (7.6)$$

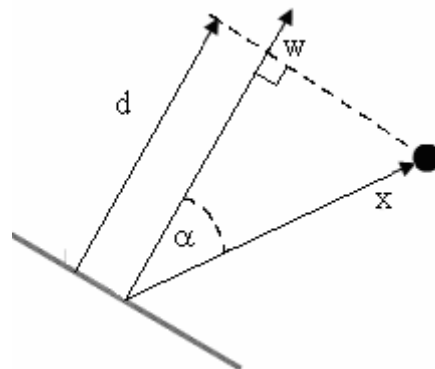
s normálou  $w$  a vzdáleností od počátku  $|-1 - b| / \|w\|$ . Nadroviny  $H_1$  a  $H_2$  mají stejné normály a neleží mezi nimi žádná data. Body  $x_i$  splňující podmínku (7.5), respektive (7.6), leží na nadrovinách  $H_1, H_2$  a nazývají se podpůrné vektory. Na obrázku (7.2) jsou tyto body označeny kroužkem [2]. Hledáme tedy takové body  $x_i$ , které splňují podmínku:

$$\min_{x_i} |w^T x + b| = 1. \quad (7.7)$$

Vzdálenost  $d(w, b; x)$  od bodu  $x$  k nadrovině  $w, b$  je definována takto:

$$\cos \alpha = \frac{w^T x_i}{\|w\| \cdot \|x_i\|}, \quad \cos \alpha = \frac{d}{\|x_i\|} \quad (7.8)$$

$$d(w, b; x) = \frac{|w^T x + b|}{\|w\|}. \quad (7.9)$$



**Obr.7.3:** Vzdálenost od bodu dané třídy k optimální nadrovině.

Optimální oddělovací nadrovina maximalizující mezeru udává velikost mezery.

$$\begin{aligned}
 \rho(w, b) &= d_+ + d_- = \min_{\{x_i; y_i=1\}} d(w, b; x_i) + \min_{\{x_i; y_i=-1\}} d(w, b; x_i) \\
 &= \min_{\{x_i; y_i=1\}} \frac{|w^T x + b|}{\|w\|} + \min_{\{x_i; y_i=-1\}} \frac{|w^T x + b|}{\|w\|} = \frac{1}{\|w\|} \left( \min_{\{x_i; y_i=1\}} |w^T x_i + b| + \min_{\{x_i; y_i=-1\}} |w^T x_i + b| \right) \\
 &= \frac{2}{\|w\|}
 \end{aligned} \tag{7.10}$$

Aby byla mezera maximální, je nutné maximalizovat výraz (7.10) neboli minimalizovat  $\|w\|$ . Pro nalezení oddělovací nadroviny je potřeba nalézt vázaný extrém, který bude funkcí  $w, b$ . Jedná se o konvexní optimalizační úlohu, která se řeší pomocí Lagrangeových funkcí. Je použito minimalizačního kritéria Lagrangeových koeficientů pro nalezení optimální oddělovací nadroviny.

$$L_p = \frac{1}{2} \|w\|^2 - \sum_{i=1}^l \alpha_i (w^T x_i + b) y_i + \sum_{i=1}^l \alpha_i \tag{7.11}$$

Takto bývá označován primární problém, kde  $\alpha_i$  jsou Lagrangeovy koeficienty. Dalším krokem algoritmu je minimalizování funkce  $L_p$  s ohledem na  $w, b$  a maximalizování s ohledem na  $\alpha_i$ . Tento problém bývá nazýván Wolfeho duální problém. Výhodou duální úlohy je to, že zde již nemusíme optimalizovat vektor, ale pouze reálná čísla  $\alpha_i$  za podstatně jednodušších omezení, než tomu bylo v primární úloze. Další velkou výhodou je fakt, že data se v duální úloze vyskytují pouze ve formě skalárních součinů [2].

$$\max_{\alpha} L_D(\alpha) = \max_{\alpha} \{ \min_{w, b} L_p \} \tag{7.12}$$

Minimum  $\min_{w, b} L_p$  určíme pomocí *Karush–Kuhn–Tucker* podmínek.

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^l \alpha_i y_i = 0 \tag{7.13}$$

$$\frac{\partial L}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^l \alpha_i y_i x_i \tag{7.14}$$

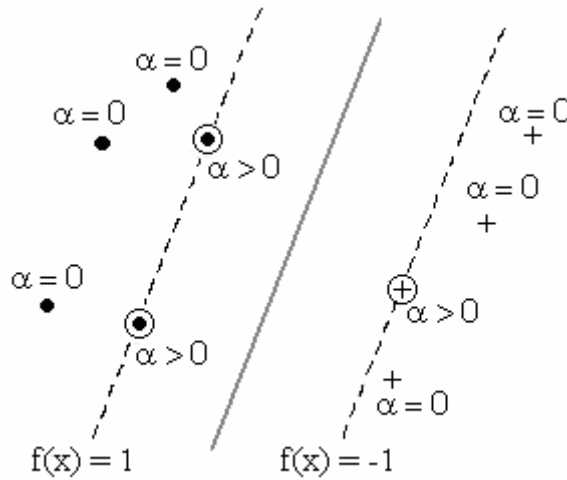
Ze vztahů (7.12), (7.13) a (7.14) plyne:

$$\max_{\alpha} L_D(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j x_i^T x_j. \quad (7.15)$$

Řešení rovnice (7.15) je:

$$\alpha_i = \arg \max_{\alpha} \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j x_i^T x_j \quad (7.16)$$

za podmínek  $\alpha_i \geq 0$  a (7.13). Vzhledem k omezením při hledání  $\alpha_i$  platí, že lineární oddělovač má nulové váhy  $\alpha_i$  pro každý datový bod kromě těch bodů, které jsou nejbližší vlastnímu oddělovači. Tyto nejbližší body se nazývají podpůrné vektory (support vectors) vzhledem k tomu, že jejich funkcí je podpora oddělovací nadroviny, která na nich spočívá. Ostatní body nejsou pro oddělovač vůbec zapotřebí, takže metoda SVM je schopna nalézt pouze ty trénovací příklady, které jsou podstatné pro oddělení jednotlivých tříd. Velkou výhodou také je, že podpůrných vektorů je obvykle mnohem méně než datových bodů [19].



**Obr.7.4:** Body ležící na nadrovinách  $H_1, H_2$  mají nenulové  $\alpha$ .

Máme tedy Lagrangeovy koeficienty  $\alpha_i$ , které nám pomohou určit parametry  $w$  a  $b$ .

$$w = \sum_{i=1}^{N_i} \alpha_i x_i y_i \quad (7.17)$$

$$y_i \cdot (w^T x_i + b) = 1 \Rightarrow b = \frac{1 - y_i w^T x_i}{y_i} \quad (7.18)$$

Rozhodovací funkci klasifikátoru pro lineární SVM lze formulovat vztahem:

$$f(x) = \sum_{i=1}^{N_s} \alpha_i y_i x_i^T x + b, \quad (7.19)$$

kde  $N_s$  je počet podpůrných vektorů.

## 7.2 Optimální separující nadrovina pro neseparabilní data

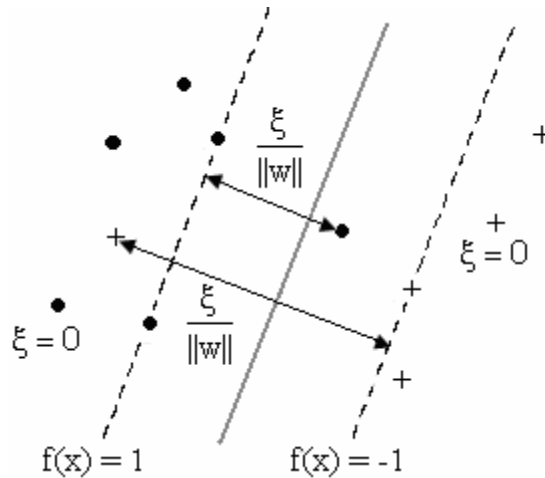
Pro neseparabilní data nenajde výše popsany algoritmus vhodné řešení. Proto je potřeba ho upravit tak, aby fungoval i na neseparabilní data. Omezení definující oddělovací nadrovinu jsou:

$$w^T x_i + b \geq +1 - \xi_i \text{ pro } y_i = +1 \quad (7.20)$$

$$w^T x_i + b \leq -1 + \xi_i \text{ pro } y_i = -1 \quad (7.21)$$

$$\xi_i \geq 0 \forall i, \quad (7.22)$$

kde  $\xi_i$  značí vzdálenost vzorku  $i$  od hranice jeho třídy, pokud se vzorek  $i$  nachází na špatné straně hranice.



**Obr.7.5:** Lineární rozdělovací nadrovina pro neseparabilní data.

Tímto je rozšířen prostor příznaků. Výraz  $\sum_i \xi_i$  definuje odhad chyby na trénovacích

datech. Místo hledání minima  $\frac{\|w\|^2}{2}$  pro separabilní případ, bude hledáno minimum:

$$\frac{\|w\|^2}{2} + C \left( \sum_i \xi_i \right)^k, \quad (7.23)$$

kde  $C$  je parametr, který si volí uživatel a odpovídá penalizaci za danou chybu. Nízká hodnota dovoluje klasifikátoru dělat chyby, to znamená, že klasifikátor dělá chyby, které ho moc nestojí (chyby chápe jako málo závažné). Zvyšující se hodnota parametru  $C$  zvyšuje cenu chybné klasifikace a vynutí si tak vytvoření přesnějšího modelu, který však nemusí být dostatečně obecný. Pro  $k=1$  je vhodné, že se  $\xi$  ani Lagrangeovy multiplikátory neobjeví ve Wolfově duálním problému maximalizace. Analogicky jako pro separabilní data transformujeme problém do duální úlohy [2].

$$\alpha_i = \arg \max_{\alpha} \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j x_i^T x_j \quad (7.24)$$

$$0 \leq \alpha_i \leq C \quad (7.25)$$

$$\sum_i \alpha_i y_i = 0 \quad (7.26)$$

Řešením je:

$$w = \sum_{i=1}^{N_s} \alpha_i x_i y_i, \quad (7.27)$$

kde  $N_s$  je počet podpůrných vektorů.

### 7.3 Nelineární SVM

Nelineární funkce je nutné použít v případě, když rozhodovací funkce není lineární. Trénovaná data jsou přetransformována z prostoru  $R^d$  do  $H$ .

$$\phi: R^d \rightarrow H \quad (7.28)$$

Řešením je tedy hledání vícerozměrného prostoru, kde jsou data separabilní a lze využít přístup pro lineární klasifikaci. Trénovací algoritmus závisí pouze na datech ze skalárního součinu v  $H$  ( $\phi(x_i) \cdot \phi(x_j)$ ). Pokud existuje Kernel funkce:

$$K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j), \quad (7.29)$$

nemusíme funkci  $\phi$  počítat, ani nemusí být známa. Ne vždy lze jednoduše určit, zda existuje pár  $\{H, \phi\}$ . Odpověď lze určit pomocí Mercerovy podmínky [2].

Nechť existuje:

$$K(x_i, y_i) = \sum_i \phi(x)_i \cdot \phi(y)_i \quad (7.30)$$

a jakákoliv funkce  $g(x)$  taková, že integrál (7.31) je konečný:

$$\int g(x)^2 dx, \quad (7.31)$$

potom pár  $\{H, \phi\}$  existuje pouze, když platí:

$$\int K(x, y) g(x)^2 g(y) dx dy \geq 0. \quad (7.32)$$

Funkce, počítající skalární součin v transformovaném prostoru, se nazývají jádrové funkce (kernel function). Například u neuronových sítí je nutné nastavovat počty skrytých vrstev či počty neuronů v jednotlivých vrstvách, přičemž podobně je tomu i u SVM. Nastavení parametrů není triviální, možný rozsah hodnot pro nastavení může být velmi velký. SVM je v terminologii neuronových sítí sít' s jednou skrytou vrstvou tvořenou jádrovými jednotkami a jednou prahovou výstupní jednotkou [12]. Rozlišujeme tři typy jádrových funkcí.

Neuronová:  $K(x_i, x_j) = \tanh(x_i^T x_j - \Theta) \quad (7.33)$

Polynomiální:  $K(x_i, x_j) = (x_i^T x_j + 1)^d \quad (7.34)$

Radiální základní funkce RBF:  $K(x_i, x_j) = \exp(-\|x_i - x_j\|^2) \quad (7.35)$

Rozhodovací funkci klasifikátoru pro nelineární SVM lze formulovat vztahem:

$$f(x) = \sum_{i=1}^{N_s} \alpha_i y_i K(s_i, x_i) + b, \quad (7.36)$$

kde  $s_i$  jsou podpůrné vektory.

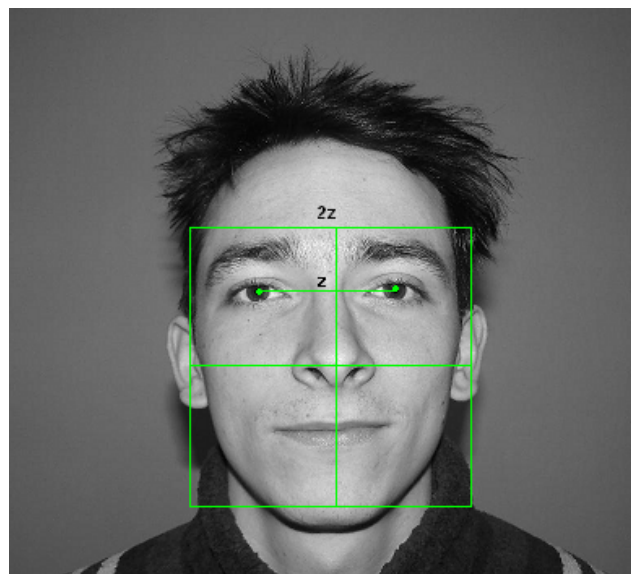


## 8. Experimenty

Byly vybrány a otestovány pouze exaktní matematické metody pro rozpoznávání lidských emocí na základě pořízeného obrazu obličeje. Neuronové sítě a podpůrné vektory SVM nebyly zařazeny do okruhu zkoumaných algoritmů. Neuronové sítě mají totiž své opodstatněné použití především při modelování složitých, často silně nelineárních soustav, kde býváme většinou postaveni před problém, že daný proces není možné s uspokojivou přesností matematicky popsat, nebo exaktní matematický model procesu je tak složitý, že jeho případná algoritmizace je buď časově a programově velmi náročná nebo dokonce nemožná. Pro účely testování jednotlivých algoritmů byla vytvořena databáze obličejů, kde se nachází 34 osob, přičemž pro každou z nich jsou k dispozici čtyři odlišné emoční výrazy (úsměv, překvapení, hněv, neutrální výraz).

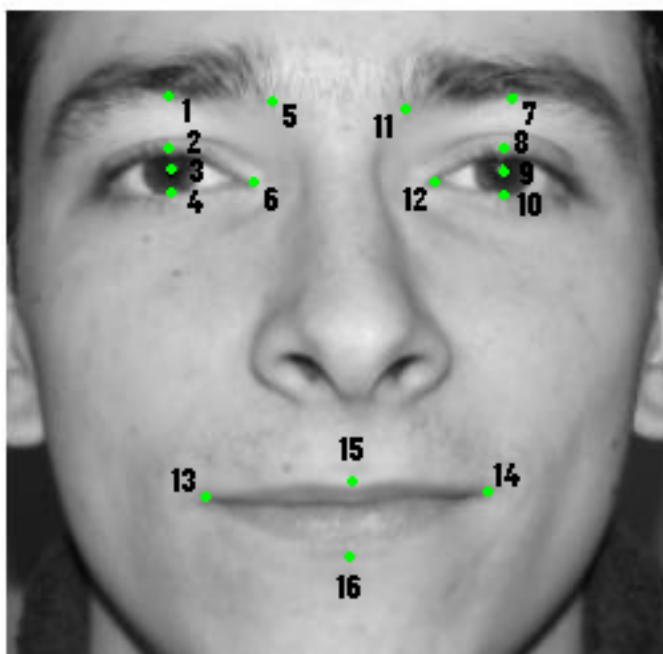
### 8.1 Popis výrazu obličeje

Ještě před vyjmutím specifických rysů je nutné, aby byly všechny obrázky normalizované, tzn. že obličeje musí mít na obrázku vždy stejnou rotaci a měřítko. Jako referenční body byly zvoleny oči, které jsou označeny na obrázku ručně. Vzdálenost manuálně získaných pozic očí je využita pro nalezení čtverce, který představuje hledanou oblast obličeje v pořízené fotografii. Takto získanou čtvercovou oblast potom normalizujeme na velikost  $300 \times 300$  pixelů.



**Obr.8.1:** Nalezení oblasti obličeje.

Jednotlivé emoční výrazy lze popsat pomocí matematického modelu, kde na základě vzdáleností mezi označenými body můžeme významně rozlišovat emoční pocity člověka, a proto je v každém normalizovaném obrázku ve vytvořené databázi ručně označeno 16 specifických bodů.



**Obr.8.2:** Šestnáct ručně označených bodů.

### Výpočet vzdáleností

$$\text{Výška očí} \quad (Y4-Y2), (Y10-Y8) \quad (8.1)$$

$$\text{Vzdálenost obočí od duhovky} \quad (Y3-Y1), (Y9-Y7) \quad (8.2)$$

$$\text{Šířka úst} \quad (X14-X13) \quad (8.3)$$

$$\text{Výška úst} \quad (Y16-Y15) \quad (8.4)$$

$$\text{Vzdálenost obočí od kraje oka} \quad (Y6-Y5), (Y12-Y11) \quad (8.5)$$

Uvedené výrazy  $X, Y$  představují souřadnice bodů ve dvourozměrném prostoru. Takto vypočtené vzdálenosti jsou dále zpracovávány algoritmy pro rozpoznávání lidských emocí, pomocí kterých bude rozhodnuto o identitě dané emoce.

## 8.2 Výsledky klasifikace

### 8.2.1 Výsledky metody PCA

Vypočtené vzdálenosti byly podrobeny metodě PCA, která redukuje dimenzi původních dat a urychluje tak následný proces identifikace. Byly zde testovány dvě metriky, na jejímž základě bylo rozhodnuto o tom, k jaké třídě emocí příslušný výraz patří.

PCA	Radost	Překvapení	Hněv	Normální	Výsledek[%]
Radost	<b>11</b>	0	0	<b>1</b>	91,6
Překvapení	0	<b>12</b>	0	0	100
Hněv	0	0	<b>12</b>	0	100
Normální	0	0	0	<b>12</b>	100

**Tab.8.1:** Klasifikační tabulka pro PCA metodu s Euklidovou metrikou. Tabulka znázorňuje, do jaké třídy výrazů je zařazen vyšetřovaný výraz. Na konci tabulky je uveden výsledek klasifikace daného výrazu v procentech (ostatní tabulky znázorňují totéž).

PCA	Radost	Překvapení	Hněv	Normální	Výsledek[%]
Radost	<b>12</b>	0	0	0	100
Překvapení	0	<b>12</b>	0	0	100
Hněv	0	0	<b>12</b>	0	100
Normální	0	0	0	<b>12</b>	100

**Tab.8.2:** Klasifikační tabulka pro PCA metodu s Mahalanobisovou metrikou.

Výsledné rozpoznávací skóre za použití Euklidovy metriky bylo 97,9%. V případě Mahalanobisovy metriky byla celková úspěšnost 100%.

### 8.2.2 Výsledky metody RCA

Druhou testovanou metodou byla RCA, kde nemusíme znát, k jaké třídě v trénované množině příslušný výraz patří. Z tabulky (8.3) je patrné, že s touto metodou bylo dosaženo za použití Euklidovy metriky nejhorších výsledků.

RCA	Radost	Překvapení	Hněv	Normální	Výsledek[%]
Radost	<b>10</b>	0	0	<b>2</b>	83,3
Překvapení	0	<b>12</b>	0	0	100
Hněv	0	0	<b>12</b>	0	100
Normální	<b>3</b>	0	0	<b>9</b>	75

**Tab.8.3:** Klasifikační tabulka pro RCA metodu s Euklidovou metrikou.

RCA	Radost	Překvapení	Hněv	Normální	Výsledek[%]
Radost	<b>12</b>	0	0	0	100
Překvapení	0	<b>12</b>	0	0	100
Hněv	0	0	<b>12</b>	0	100
Normální	0	0	0	<b>12</b>	100

**Tab.8.4:** Klasifikační tabulka pro RCA metodu s Mahalanobisovou metrikou.

Celkové rozpoznávací skóre za použití Euklidovy metriky bylo 89,6%. V případě Mahalanobisovy metriky bylo celkového rozpoznávací skóre 100%.

### 8.2.3 Výsledky metody IPCA

Další testovanou metodou byla IPCA, kde je charakteristický prostor reprezentující jednotlivé emoce vytvářen pro každý výraz zvlášť, a proto jsou dobře potlačeny společné charakteristické rysy jednotlivých tříd. Tato skutečnost pak vede k tomu, že je metoda IPCA odolná vůči vnitřní variabilitě. Nutnou podmínkou ovšem je, že musíme v trénované množině dat znát, k jaké třídě jednotlivé výrazy patří.

IPCA	Radost	Překvapení	Hněv	Normální	Výsledek[%]
Radost	<b>12</b>	0	0	0	100
Překvapení	0	<b>12</b>	0	0	100
Hněv	0	0	<b>12</b>	0	100
Normální	0	0	0	<b>12</b>	100

**Tab.8.5:** Klasifikační tabulka pro IPCA metodu.

### 8.2.4 Výsledky metody FLDA

Principem FLDA je nalézt takovou projekci dat do prostoru nižší dimenze, která umožňuje co nejzřetelnější oddělení jednotlivých tříd. Nutnou podmínkou ovšem je, že musíme v trénované množině znát, k jaké třídě jednotlivé výrazy patří.

FLDA	Radost	Překvapení	Hněv	Normální	Výsledek[%]
Radost	<b>11</b>	0	0	<b>1</b>	91,6
Překvapení	0	<b>12</b>	0	0	100
Hněv	0	0	<b>12</b>	0	100
Normální	0	0	0	<b>12</b>	100

**Tab.8.6:** Klasifikační tabulka pro FLDA metodu s Euklidovou metrikou.

<b>FLDA</b>	Radost	Překvapení	Hněv	Normální	Výsledek[%]
Radost	<b>12</b>	0	0	0	100
Překvapení	0	<b>12</b>	0	0	100
Hněv	0	0	<b>12</b>	0	100
Normální	0	0	0	<b>12</b>	100

**Tab.8.7:** Klasifikační tabulka pro FLDA metodu s Mahalanobisovou metrikou.

Celkové rozpoznávací skóre za použití Euklidovy metriky bylo 97,9%. V případě Mahalanobisovy metriky bylo celkové rozpoznávací skóre 100%.

### 8.2.5 Výsledky metody LDA

Metoda LDA využívá pro správnou klasifikaci jednotlivých emocí do příslušných tříd podmíněné hustoty pravděpodobností, které odpovídají normálnímu rozložení. S touto metodou bylo dosaženo celkové úspěšnosti 100%.

<b>LDA</b>	Radost	Překvapení	Hněv	Normální	Výsledek[%]
Radost	<b>12</b>	0	0	0	100
Překvapení	0	<b>12</b>	0	0	100
Hněv	0	0	<b>12</b>	0	100
Normální	0	0	0	<b>12</b>	100

**Tab.8.8:** Klasifikační tabulka pro metodu LDA.

### 8.2.6 Výsledky metody FLDA+RCA

Přístup posledního testovaného algoritmu je založen na spojení RCA a FLDA. Pro nalezení prostoru, jenž bude vhodnou reprezentací dané emoce, se využívá FLDA a data promítnutá do tohoto prostoru jsou podrobena metodě RCA (Mahalanobisova metrika).

<b>FLDA+RCA</b>	Radost	Překvapení	Hněv	Normální	Výsledek[%]
Radost	<b>12</b>	0	0	0	100
Překvapení	0	<b>12</b>	0	0	100
Hněv	0	0	<b>12</b>	0	100
Normální	0	0	0	<b>12</b>	100

**Tab.8.9:** Klasifikační tabulka pro metodu FLDA+RCA.

Výsledné rozpoznávací skóre bylo 100%.

## Závěr

V rámci diplomové práce byl vytvořen systém pro rozpoznávání lidských emocí na základě pořízeného obrazu obličeje. Bylo otestováno celkem 5 algoritmů založených na strukturálních metodách rozpoznávání. Práce obsahuje rešerši na dané téma, kde jsou prezentovány jednotlivé přístupy pro rozpoznávání emocí, ze kterých vytvořené algoritmy vycházejí. Z popisovaných studií v druhé části práce je patrné, že nejlepší výsledky byly dosaženy pomocí neuronových sítí, které ovšem nejsou zahrnuty v žádném z testovaných algoritmů. Neuronové sítě mají totiž své opodstatněné použití především v případech, kdy při řešení daného problému není možné matematicky popsat všechny vztahy a souvislosti, které ovlivňují sledovaný proces, protože exaktní matematický model procesu je tak složitý, že jeho případná algoritmizace je časově a programově náročná, v některých případech dokonce nemožná, nebo pokud s žádnou jinou metodou nelze dosáhnout uspokojivého výsledku.

Vybrané metody byly otestovány na vytvořené databázi obličejů, kde se nachází 34 osob, přičemž pro každou z nich jsou k dispozici čtyři odlišné emoční výrazy (úsměv, překvapení, zlost, neutrální výraz). Nejlepšího rozpoznávacího skóre 100% dosahovalo hned několik exaktních matematických metod. Ze dvou testovaných metrik vykazovala vyšší úspěšnost Mahalanobisova metrika. U všech algoritmů, kde byla použita tato metrika, bylo dosaženo vyššího rozpoznávacího skóre než s Euklidovou metrikou. Nejhorší výsledek vykazovala metoda RCA s Euklidovou metrikou, úspěšnost byla pouze 89,6%.

Oproti jiným oborům informatiky je odvětví zabývající se rozpoznáváním emocí počítačem méně formalizované a přístupy autorů se velmi liší. Z práce tedy vyplývá, že je problematika automatického rozpoznávání emocí podle obličeje neobyčejně rozsáhlá a složitá, a proto nabízí mnoho prostoru pro vylepšování stávajících a vývoj nových algoritmů.

## Použitá literatura

- [1] Belhumeur, Peter N. – Hespanha, Joao P. – Kriegman, David J. *Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection*. [online]. [cit. 2009-2-25]. URL: <<http://www1.cs.columbia.edu/~belhumeur/journal/fisherface-pami97.pdf>>
- [2] Burges, Christopher J.C. *A Tutorial on Support Vector Machines for Pattern Recognition*. [online]. [cit. 2009-4-3]. URL: <<http://www.umi.acs.umd.edu/~joseph/support-vector-machines4.pdf>>
- [3] Císař, P. *Využití metod odezírání ze rtů pro podporu rozpoznávání řeči*. Plzeň, 2004. Disertační práce. Fakulta aplikovaných věd. Západočeská univerzita v Plzni. [online]. [cit. 2008-12-28]. URL: <[http://ui.zcu.cz/img/p\\_249.pdf](http://ui.zcu.cz/img/p_249.pdf)>
- [4] Daw-Tung, Lin. *Facial Expression Classification Using PCA and Hierarchical Radial Basis Function Network*. [online]. [cit. 2008-12-7]. URL: <[http://www.iis.sinica.edu.tw/page/jise/2006/200609\\_03.pdf](http://www.iis.sinica.edu.tw/page/jise/2006/200609_03.pdf)>
- [5] Fatka, O. *Předpovědi hydrogramů průtokových vln v říčním systému s využitím neuronových sítí*. Výzkum na Fakultě stavební ČVUT v Praze. [online]. [cit. 2009-1-3]. URL: <[http://hydraulika.fsv.cvut.cz/vyzkum/nejistoty/predpoved/Nejistoty\\_internet2.pdf](http://hydraulika.fsv.cvut.cz/vyzkum/nejistoty/predpoved/Nejistoty_internet2.pdf)>
- [6] Havlík, J. *Klasifikace pohybů části lidského těla v prostoru*. Praha, 2008. Disertační práce. FEL ČVUT v Praze. [online]. [cit. 2009-2-11]. URL: <[http://amber.feld.cvut.cz/user/xhavlikj/documents/2008\\_01.pdf](http://amber.feld.cvut.cz/user/xhavlikj/documents/2008_01.pdf)>
- [7] Hertz, T. – Pavel, M. – Shental, N. – Weinshall, D. *Adjustment Learning and Relevant Component Analysis*. [online]. [cit. 2008-12-27]. URL: <[www.cs.huji.ac.il/~daphna/papers/rca-eccv.ps.gz](http://www.cs.huji.ac.il/~daphna/papers/rca-eccv.ps.gz)>
- [8] Hlaváč, V. *Support Vector Machines*. [online]. [cit. 2009-4-5]. URL: <<http://cmp.felk.cvut.cz/~hlavac/TeachPresEn/31PattRecog/32-06SVM.pdf>>
- [9] Changing Minds. *Basic emotions*. [online]. [cit. 2008-11-18]. URL: <<http://changingminds.org/explanations/emotions/basic%20emotions.htm>>
- [10] Jarolímková, H. *Počítačové zpracování emocí*. Brno, 2007. Diplomová práce na Fakultě informatiky Masarykovy univerzity v Brně. [online]. [cit. 2008-11-18]. URL: <[http://is.muni.cz/th/60784/fi\\_m/60784\\_diplomova\\_prace.pdf](http://is.muni.cz/th/60784/fi_m/60784_diplomova_prace.pdf)>

- [11] Khanmohammadi, S. – Seyedarabi, H. – Aghagolzadeh, A. *Facial Expressions Animation and Lip Tracking Using Facial Characteristic Points and Deformable Model*. [online]. [cit. 2009-1-19]. URL: <<http://www.waset.org/ijit/v1/v1-4-37.pdf>>
- [12] Kůrková, V. *Strojové učení se schopností generalizace*. Ústav informatiky. Akademie věd České republiky. [online]. [cit. 2009-4-8]. URL: <<http://hilbert.chtf.stuba.sk/KUZV/download/kuzv-kurkova.pdf>>
- [13] Liu X. – Chen, T. – Kumar, B.V.K.V. *Face Authentication for Multiple Subjects Using Eigenflow*. [online]. [cit. 2009-1-15]. URL: <[http://amp.ece.cmu.edu/Publication/Xiaoming/PR\\_eigenflow\\_xiaoming.pdf](http://amp.ece.cmu.edu/Publication/Xiaoming/PR_eigenflow_xiaoming.pdf)>
- [14] Paleček, K. *Rozpoznávání osob na základě detekovaného obličeje*. Liberec, 2007. Bakalářská práce. Fakulta mechatroniky a mezioborových inženýrských studií TUL. [cit. 2009-1-10].
- [15] Polikar, R. *Pattern Recognition*. [online]. [cit. 2009-2-27]. URL: <<http://users.rowan.edu/~polikar/RESEARCH/PUBLICATIONS/wiley06.pdf>>
- [16] Shakhnarovich, G. – Moghaddam, B. *Face Recognition in Subspaces*. [online]. [cit. 2009-1-10]. URL: <<http://www.merl.com/reports/docs/TR2004-041.pdf>>
- [17] Shen-Chuan, T. - Hung-Fu, J. - Kuo-Chen, Ch. - Yu-Yi, L. - Chien-Shiang, H. *Facial Expression Recognition using Neural Network*. [online]. [cit. 2009-1-19]. URL: <<http://isis2007.fuzzy.or.kr/submission/upload/a1189.pdf>>
- [18] Sorci, M. – Antonini, G. – Thiran, Jean-Philippe. *Fisher's Diskriminant and Relevant Component Analysis for Static Facial Expression Classification*. [online]. [cit. 2008-12-27]. URL: <[http://infoscience.epfl.ch/record/87190/files/Sorci2005\\_1414.pdf](http://infoscience.epfl.ch/record/87190/files/Sorci2005_1414.pdf)>
- [19] *Support vector machines (SVM)*. [online]. [cit. 2009-3-25]. URL: <[http://is.muni.cz/el/1433/podzim2006/PA034/09\\_SVM.pdf?fakulta=1433;obdobi=3523;kod=PA034](http://is.muni.cz/el/1433/podzim2006/PA034/09_SVM.pdf?fakulta=1433;obdobi=3523;kod=PA034)>
- [20] Šembera, J. *Aplikace analýzy hlavních komponent pro redukci dimenze transportě-reakčního problému*. [online]. [cit. 2009-1-10]. URL: <[http://math.fce.vutbr.cz/~pribyl/workshop\\_2007/prispevky/Sembera.pdf](http://math.fce.vutbr.cz/~pribyl/workshop_2007/prispevky/Sembera.pdf)>
- [21] Štuková, J. *Neuronové sítě v řízení systémů*. [online]. [cit. 2008-11-18]. URL: <<http://www.fs.vsb.cz/books/NeuronoveSite/>>



- [22] Tailor, C.J. - Cootes T.F. *Statistical Models of Appearance for Computer Vision* [online]. [cit. 2009-5-7]. URL: <[http://www.face-rec.org/algorithms/AAM/app\\_models.pdf](http://www.face-rec.org/algorithms/AAM/app_models.pdf)>
- [23] Teknomo, K. *Discriminant Analysis Tutorial*. [online]. [cit. 2009-2-10]. URL: <<http://people.revoledu.com/kardi/tutorial/LDA/index.html>>
- [24] Trias, M. *Face verification based on Support Vector Machines*. [online]. [cit. 2008-11-26]. URL: <<http://infoscience.epfl.ch/getfile.py?mode=best&recid=87169>>
- [25] Trojanová, J. *Rozpoznávání emocí z výrazu v obličeji*. [online]. [cit. 2008-11-18]. URL: <<http://musslap.zcu.cz/cs/rozpovnavani-emoci/>>
- [26] Veselovský, M. *Neuronové sítě*. [online]. [cit. 2009-3-18]. URL: <<http://www.avari.cz/uir/index.php?pq=vsechno>>
- [27] Visutsak, P. *Emotion Recognition through Lower Facial Expressions Using Support Vector Machines*. [online]. [cit. 2008-12-16]. URL: <<http://www.kmitl.ac.th/~s0067152/PDF/Emotion%20Recognition.pdf>>
- [28] Wikipedia. *Paul Ekman*. [online]. [cit. 2008-11-18]. URL: <[http://en.wikipedia.org/wiki/Paul\\_Ekman](http://en.wikipedia.org/wiki/Paul_Ekman)>